

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2002, 15:19:20 ; Search time 15 Seconds

(without alignments)
2514.189 Million cell updates/sec

Title: US-09-744-125-2

Perfect score: 5225
Sequence: 1 MALRGVRYKLVCAEKNDAK.....RKSLCHQPHRRPCPNR 974

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5225	100.0	1001	1	TP3A_HUMAN
2	4586	87.8	1003	1	TP3A_MOUSE
3	2230.5	42.7	1250	1	TP3A_MOUSE
4	1711	32.7	759	1	TP3A_MOUSE
5	1279	24.5	622	1	TP3A_MOUSE
6	1174.5	22.5	656	1	TP3A_MOUSE
7	1078	20.6	862	1	TP3A_MOUSE
8	1066	20.4	862	1	TP3A_MOUSE
9	1022	19.6	875	1	TP3A_MOUSE
10	688	13.2	663	1	TP3A_MOUSE
11	653	12.5	770	1	TP3A_MOUSE
12	600.5	11.5	685	1	TP3A_MOUSE
13	591	11.0	686	1	TP3A_MOUSE
14	571	11.0	761	1	TP3A_MOUSE
15	563	10.8	718	1	TP3A_MOUSE
16	483.5	9.3	710	1	TP3A_MOUSE
17	467	8.9	673	1	TP3A_MOUSE
18	449.5	8.6	700	1	TP3A_MOUSE
19	441	8.4	690	1	TP3A_MOUSE
20	441	8.4	876	1	TP3A_MOUSE
21	426	8.2	691	1	TP3A_MOUSE
22	420	8.0	865	1	TP3A_MOUSE
23	413	7.9	865	1	TP3A_MOUSE
24	412	7.9	1217	1	TP3A_MOUSE
25	406	7.8	868	1	TP3A_MOUSE
26	404	7.7	868	1	TP3A_MOUSE
27	403.5	7.7	868	1	TP3A_MOUSE
28	398	7.6	868	1	TP3A_MOUSE
29	396.5	7.6	868	1	TP3A_MOUSE
30	382	7.3	1060	1	TP3A_MOUSE
31	379.5	7.3	776	1	TP3A_MOUSE
32	377.5	7.2	736	1	TP3A_MOUSE
33	375.5	7.2	776	1	TP3A_MOUSE

34	370.5	7.1	952	1	TOP1_STRCO	09x909 streptomyc
35	370	7.1	660	1	TOP1_XYIFA	09phk2 xylicella fas
36	367	7.0	731	1	TOP1_TREPA	083409 treponema p
37	359.5	6.9	736	1	TOP1_HELPJ	09zmv7 helicobacte
38	356.5	6.8	861	1	TOP1_BUCAI	057371 buchnera ap
39	349	6.7	540	1	TOP1_AOUAE	066893 aquifex aco
40	344.5	6.6	650	1	TOP1_PASMU	09cp53 pasteurella
41	339	6.5	653	1	TOP1_ECOLI	09cp53 escherichia
42	336.5	6.4	711	1	TOP1_MYCPN	09cp53 mycoplasma
43	329	6.3	649	1	TOP1_SALTY	09kgf5 vibrio chol
44	328	6.3	647	1	TOP1_VIBCH	09kgf5 vibrio chol
45	325.5	6.2	651	1	TOP1_HAEIN	09kgf5 haemophilus

ALIGNMENTS

```

RESULT 1
TP3A_HUMAN
ID TP3A_HUMAN STANDARD: PRT: 1001 AA.
AC Q13472: Q13473:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA topoisomerase III alpha (EC 5.99.1.2).
GN TOP3A OR TOP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96195027; PubMed=8622991;
RA Hanai R., Caron P.R., Wang J.C.;
RT Human TOP3: a single-copy gene encoding DNA topoisomerase III.;
RL Proc. Natl. Acad. Sci. U.S.A. 93:3653-3657(1996).
CC -!- FUNCTION: REDUCES THE NUMBER OF SUPERCOILS IN A HIGHLY NEGATIVELY
CC SUPERCOILED DNA.
CC -!- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC DNA, followed by passage and rejoining.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: HIGH EXPRESSION IS FOUND IN TESTIS, HEART,
CC SKELETAL MUSCLE AND PANCREAS.
CC -!- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U43431; AAB03694.1; -
CC EMBL: U43431; AAB03695.1; -
CC MIM: 601243; -
CC InterPro: IPR003601; DNATOP1_ATP_bind.
CC InterPro: IPR003602; DNATOP1_DNA_bind.
CC InterPro: IPR000380; Pro_topoisomerase.
CC InterPro: IPR002936; Toprim.
CC InterPro: IPR001878; Znf_CCHC.
CC Pfam: PF01131; Topoisom_bac; 1.
CC Pfam: PF01751; Toprim; 1.
CC Pfam: PF01396; zf-C4_Topoisom; 1.
CC Pfam: PF00098; zf-CCHC; 1.
CC PRINTS: PR00939; C2HCNFRNGER.
CC PRINTS: PR00417; PTPISMRASET.
CC SMART: SM00437; TOP1AC; 1.
CC SMART: SM00436; TOP1BC; 1.
CC SMART: SM00493; TOPRIM; 1.
CC SMART: SM00343; Znf_C2HC; 1.

```

DR PROSITE: PS00396; TOPOISOMERASE_I_PROK; 1.
 KM Isomerase; Topoisomerase; DNA-binding; Repeat; Zinc-finger;
 KW Alternative splicing; Polymorphism.
 FT ACT_SITE 362 362 DNA CLEAVAGE (BY SIMILARITY).
 FT DOMAIN 658 685 C4-TYPE (POTENTIAL).
 FT DOMAIN 812 923 2 x 27 AA APPROXIMATE REPEATS.
 FT REPEAT 812 839 1.
 FT REPEAT 896 923 2.
 FT VARSPLIC 1 25 MISSING (IN SHORT ISOFORM).
 FT VARIANT 596 596 C->Y.
 SQ SEQUENCE 1001 AA; 112372 MW; 06558C749569ECC2 CRC64;

Query Match 100.0%; Score 5225; DB 1; Length 1001;
 Best Local Similarity 100.0%; Pired. No. 0;
 Matches 974; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALGVRVLCVAREKNDKAGIADLLSNGRRRRREGSKFKIYEPDQHLGQNTVMYT 60
 DB 28 MALGVRVLCVAREKNDKAGIADLLSNGRRRRREGSKFKIYEPDQHLGQNTVMYT 87
 QY 61 SVSGHLAHDFQOMQFRKQSCNPLVFEAEIEKYCPENFVDIKITLERETQCALVTMT 120
 DB 88 SVSGHLAHDFQOMQFRKQSCNPLVFEAEIEKYCPENFVDIKITLERETQCALVTMT 147
 QY 121 DCDREGENIGEIIHVCAPVNLQVLRARSEITPAHVRACENLTPEDDRVSDAVDR 180
 DB 148 DCDREGENIGEIIHVCAPVNLQVLRARSEITPAHVRACENLTPEDDRVSDAVDR 207
 QY 181 QELDLRIGAATRFOTLRQIFPEVLAEOILISGSCQPLGPFVVEFKIAQAVPEITF 240
 DB 208 QELDLRIGAATRFOTLRQIFPEVLAEOILISGSCQPLGPFVVEFKIAQAVPEITF 267
 QY 241 HRIKVTDHDKDIYEFNKKRRHLEFNHTACLVLYQLCVEDPMATVVEVSKSRKRPQAL 300
 DB 268 HRIKVTDHDKDIYEFNKKRRHLEFNHTACLVLYQLCVEDPMATVVEVSKSRKRPQAL 327
 QY 301 DTVELEKIAKRLINAKETRIAEKLYTQGYISTPRTETNIFPRDLMLTVLVEQOTDP 360
 DB 328 DTVELEKIAKRLINAKETRIAEKLYTQGYISTPRTETNIFPRDLMLTVLVEQOTDP 387
 QY 361 RWGFAOSILIERGGPTPRNGKSDOAHPIHPRTYTNLQDEORLYEFIRHFLACSSO 420
 DB 388 RWGFAOSILIERGGPTPRNGKSDOAHPIHPRTYTNLQDEORLYEFIRHFLACSSO 447
 QY 421 DAOQETVEIDIAQERFVAGHLMILARNYLDVYPYDHMSDKILPVYEGSGHFGPSTVEM 480
 DB 448 DAOQETVEIDIAQERFVAGHLMILARNYLDVYPYDHMSDKILPVYEGSGHFGPSTVEM 507
 QY 481 VDGETSPKLLTEADLIAEMKHGIGDATIAEHITTKAAMYGLTPDKRFLPEGLGNG 540
 DB 508 VDGETSPKLLTEADLIAEMKHGIGDATIAEHITTKAAMYGLTPDKRFLPEGLGNG 567
 QY 541 LVEGYDSMGYEMKPDIAELADKLICDGKKDFVYLROOVOKYKOVFLFEAAKAKKL 600
 DB 568 LVEGYDSMGYEMKPDIAELADKLICDGKKDFVYLROOVOKYKOVFLFEAAKAKKL 627
 QY 601 DEALAQYFGNGTELAQOEDIYPAPEPIRKCPOCKNDKAVLTKKNGGYLSGSGFPEERS 660
 DB 628 DEALAQYFGNGTELAQOEDIYPAPEPIRKCPOCKNDKAVLTKKNGGYLSGSGFPEERS 687
 QY 661 AVLPLDSVLEASRDSSVCQPHVYRLKIKFRGSLPTPTMLEFYVCCIGGCDTLREI 720
 DB 688 AVLPLDSVLEASRDSSVCQPHVYRLKIKFRGSLPTPTMLEFYVCCIGGCDTLREI 747
 QY 721 LDIRFSGGPPRASQPSRLQANOSLRMDNSHOPADSRQTSKSAQAQTLPPPTAGE 780
 DB 748 LDIRFSGGPPRASQPSRLQANOSLRMDNSHOPADSRQTSKSAQAQTLPPPTAGE 807
 QY 781 SNSVTCMGQEAVALYRKEGPNRGROFKCNGSGCNFELWADSPNPGAGPPALAYRPL 840
 DB 808 SNSVTCMGQEAVALYRKEGPNRGROFKCNGSGCNFELWADSPNPGAGPPALAYRPL 867

QY 841 GASLCGCPGPIHLCGFGNPGDGGSSGTSCLCSQPSVTRFYQKDPKNGROFHTCAKPRE 900
 DB 868 GASLCGCPGPIHLCGFGNPGDGGSSGTSCLCSQPSVTRFYQKDPKNGROFHTCAKPRE 927
 QY 901 QOCGFQWDEMTAGTSGAPSWTGDRTGTESEARSRPRASSSDMGSTAKPKKCSIC 960
 DB 928 QOCGFQWDEMTAGTSGAPSWTGDRTGTESEARSRPRASSSDMGSTAKPKKCSIC 987
 QY 961 HQPGHTRFPCQNR 974
 DB 988 HQPGHTRFPCQNR 1001

RESULT 2
 TP3A_MOUSE
 ID TP3A_MOUSE STANDARD; PRT; 1003 AA.
 AC 070157:
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA topoisomerase III alpha (BC 5.99.1.2).
 GN TP3A OR Top3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/C; TISSUE=Testis;
 RX MEDLINE=98201702; PubMed=9540825.
 RA Seki T., Seki M., Katada T., Enomoto T.;
 RT "Isolation of a cDNA encoding mouse DNA topoisomerase III which is
 RT highly expressed at the mRNA level in the testis";
 RL Biochim. Biophys. Acta 1396:127-131(1998).
 CC -1- FUNCTION: REDUCES THE NUMBER OF SUPERCOILS IN A HIGHLY NEGATIVELY
 CC SUPERCOILED DNA.
 CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
 CC DNA, followed by passage and rejoining.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS.
 CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB006074; BAA25662.1; -.
 DR MGD: MGI:1197527; Top3a.
 DR InterPro: IPR003601; DNAtopI_ATP_bind.
 DR InterPro: IPR003602; DNAtopI_DNA_bind.
 DR InterPro: IPR000380; Pro_topoisomerase.
 DR InterPro: IPR002936; Toprim.
 DR InterPro: IPR001878; Znf_CCHC.
 DR Pfam: PF01131; Topoisom_bac; 1.
 DR Pfam: PF01751; Toprim; 1.
 DR Pfam: PF01396; zf-C4_Topoisom; 1.
 DR Pfam: PF00098; zf-CCHC; 1.
 DR PRINTS: PRO0939; C2HCZNFINGER.
 DR PRINTS: PRO0417; PRTPISMRASET.
 DR SMART: SM00437; TOP1AC; 1.
 DR SMART: SM00436; TOP1BC; 1.
 DR SMART: SM00493; TOPRIM; 1.
 DR SMART: SM00343; Znf_C2HC; 1.
 DR PROSITE: PS00396; TOPOISOMERASE_I_PROK; 1.
 KW Isomerase; Topoisomerase; DNA-binding; Repeat; Zinc-finger.
 FT ACT_SITE 362 362 DNA CLEAVAGE (BY SIMILARITY).
 FT DOMAIN 658 685 C4-TYPE (POTENTIAL).
 FT DOMAIN 814 925 2 x 27 AA APPROXIMATE REPEATS.
 FT REPEAT 814 841 1.


```

CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AF255733: AAF71288.1: -.
DR      EMBL: AF203663: AAF53813.1: ALT_SEQ.
DR      FlyBase: FBgn00402568: Top3-alpha.
DR      InterPro: IPR003601: DNATop1_ATP_bind.
DR      InterPro: IPR003602: DNATop1_DNA_bind.
DR      InterPro: IPR000380: Pro_topoisomerase.
DR      InterPro: IPR002936: Toprim.
DR      InterPro: IPR001878: znf_CCHC.
DR      Pfam: PF01131: Topoisom_bac; 1.
DR      Pfam: PF01751: Toprim; 1.
DR      Pfam: PF01396: zf-C4_Topoisom; 1.
DR      PRINTS: PR00417: PRTPISMRASE1.
DR      SMART: SM00437: Top1ac; 1.
DR      SMART: SM00436: Top1bc; 1.
DR      SMART: SM00493: TOPRIM; 1.
DR      SMART: SM00343: znf_C2HC; 1.
DR      SMART: SM00396: TOPOISOMERASE_1_PROK; 1.
DR      PROSITE: PS00396: TOPOISOMERASE_1_PROK; 1.
DR      Isomerase: Topoisomerase; DNA-binding.
FT      ACT_SITE 356 356 DNA_CLEAVAGE (BY SIMILARITY).
SQ      SEQUENCE 1250 AA; 136137 MW; 288096770B3DBE75E CRC64;

Query Match 42.7%; Score 2230.5; DB 1; length 1250;
Best Local Similarity 41.6%; Pred. No. 2,1e-136;
Matches 514; Conservative 125; Mismatches 307; Indels 291; Gaps 30;

QY      8 KVLVAEKNAKAGIADLISNGMRREGSKPKIYEDVHLGQANTMTWTSVSHLL 67
DB      27 KVLVAEKNAKAGIADLISNGMAAREGYSVYNKVFEPFAPVKQAKMWTISVGHMM 86
QY      68 AHDFOMQKRWQSCNPVLEFAEIEKYCPENFVDIKKTLERETROCALVITDCEGE 127
DB      87 QLAQVSYKMRWTVDPNRSFLDAPVEKGVSDYEPKIKRTLEBEVGCQGLIWTDCDRGE 146
QY      128 NIGEITIHCAKAVPNLOVLARSEITPAVTRACENLTPEPDQVSAVVRQELDIRI 187
DB      147 NIGEITIDVCAIKPNISVYRATSEITTVAVRRALQQLGQDPKQSDPAVVRTELDIRT 206
QY      188 GAATFROTLELORIFPEVLAEOILSYGSCOPTLGAFVVERFKAIQAVPELPHRIKVTN 247
DB      207 GAATFROTMLORLFEKTIADKLISYSCOITPLGFAERKELEAFVSEPFKIKVLH 266
QY      248 DHKGIVEFNMRHRLFNHTACLVLYOLCV--EDPMATVEVRSKPKSKMRPOALDVEL 305
DB      267 TIDDLTYEFNMRARRLFPEKCECENYLLCLAEPPRALVESVTVKPKKMRPTLDIVEM 326
QY      306 EKLSAKRLINAKETMRIAEKLTYQGYISTPRTETNIFRDLNLTVLVEQOTPPRWCAGF 365
DB      327 EKLSAKRLIKSAKETMTIAEKLTYKGFISTPRTETNIFSKEFALAPLVEWOTGHRDWCAF 386
QY      366 AQSILTEGGPPRRNGKSDQAHPIHPKTYNNLQGDQORLYEFVIRFELACSSDQADGQ 425
DB      387 AQRVIE-WGPPRRNGKSDQAHPIHPKTIENLQGNARARYELVIRFELACVSDADAGS 445
QY      426 ETVVEIDIAQERFVAGHMLIARNYLDVYRYDHSMDKLPLVYEGOSHPOSTVEAVDGET 485
DB      446 ETVLVHIDIAQERFVAGHMLIARNYLDVYDYSKSAQIHHYENGQREPREVSLHEAT 505
QY      486 SPEPLLEADIALMEKHGISTDAITAHNHTITAKRMVGLTPDQKRL-PEHLSGGLVEG 544
DB      506 TAPLLEADIALMEKHGISTDAITAHNHTITAKRMVGLTPDQKRL-PEHLSGGLVEG 564
QY      545 YDSMGYEMSKPDLRAELADKLICDGKKDFVLRQOVQYKQFTEAFAVAKAKLDEAL 604

```

```

DB      565 YDAMELAIAPOLRAEFELDKLICGQKDKPVULTEDIAKKAQVQOITDKITMDAKI 624
QY      605 AQYFGN--GTLEAOQEDIIYPAMP-----EPIKCPQCNK-DMYLTKRN-GGFYLSMGF 655
DB      625 SARIEFPAANSVAQEGADGASPSHGIIQSIQPCPKCNDAFLALPKKNOQGWYICNNF 684
QY      656 PECSAVMLPDSVLEASRDSVYCPQHPHYRLKLKTRKGLP-----PMPLEPFC 709
DB      685 PDCKAVALPECKKASVLYDECCPCGCG--YRM-LKF-RLSTPYRGVFGTSGMYKTC 740
QY      710 IGCDDTLREILDLP-----SGGP----- 729
DB      741 L-PCDNLFRFTTNINDSVKKGIVGEYRGGGGPGGPGGSGRAGSGGSSGCG 799
QY      730 ----- 729
DB      800 GSGSGGSGSGSGSGTGSGGSGMGSGGGLGCGGKKPKGEGSKSATKPKPNEPKPK 859
QY      730 ----PRA-----SQPSGRLOA-----NQSINRMD----- 749
DB      860 KTKPEKAPNKKTSKSSGSIKSFPTSAAPNLSASNGLDEFFDSNDGFEDMLAAESVE 919
QY      750 NSOHPO-----PAD----- 758
DB      920 SSSQPKTISWPLDDIDIAAFADDDAEFEALVNGCTMPTESNGDQQLDKSLSEWIKED 979
QY      759 -----SKOTGSSKALQTL-----PTAAGESNSVYC-NGC 789
DB      980 KADERPMLGTRERASLCTAAPTTPPKAPRMRDVSVERDSTPPSSVPESEVLTCTGCG 1039
QY      790 QEAVLLTYRKESGPNRGROFFFC-NGSCNFMADSPNPGAGGPALAYRPLG----ASL 844
DB      1040 QPARNTVYRKNGPNGLRYKCPKPDCEPNFOWADEPPSAKSKNSTGSAOSTTSMGNS 1099
QY      845 GCPGPGIHLG-----GFGNPGDGSSTGSCSLCSQP 875
DB      1100 RVTYLPSTIQOSNSORGSMSNSSTVITITQTKKOQERNATATPGD--GGEWKNCCQL 1157
QY      876 SVTRTVQKDGPNKGRQFTCAKPREQCGFQWVDENTAPGTGAPSWTGRGRTLESEA 935
DB      1158 ASQLTRKRDGPNQGRPFYAC--PTREKSCGFPKMGEDDONOGAS--STSW-----GSANRMP 1211
QY      936 RSKPRASSDMGSTATKPKKCSLCHQPHHTRPCCPQ 972
DB      1212 GRSQPTALTSD---GPKTRCGLCRKEGHTRNCKPR 1244

RESULT 4
TOP3_CAEEL
ID TOP3_CAEEL STANDARD: PRT: 759 AA.
AC O61660;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA topoisomerase III (EC 5.99.1.2).
GN TOP3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Kim Y.-C., Koo H.-S.;
RT "cDNA cloning and overexpression of Caenorhabditis elegans DNA
RT topoisomerase III";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC "cDNA cloning and overexpression of Caenorhabditis elegans DNA
CC DNA, followed by passage and rejoining.
CC -I- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF057032: AAC13567.1: -
 DR InterPro: IPR003601: DNATopI_ATP_bind.
 DR InterPro: IPR003602: DNATopI_DNA_bind.
 DR InterPro: IPR000380: Pro_topoisomerase.
 DR InterPro: IPR002936: Toprim.
 DR Pfam: PF01131: Topoisom. bac. 1.
 DR Pfam: PF01751: Toprim. 1.
 DR PRINTS: PR00417: PRTPISMRASE1.
 DR SMART: SM00437: TOP1AC. 1.
 DR SMART: SM00436: TOP1BC. 1.
 DR SMART: SM00493: TOP1RM. 1.
 DR PROSITE: PS00396: TOPOISOMERASE_I_PROK. 1.
 DR Isomerase: Topoisomerase: DNA-binding.
 FT ACT_SITE 334 334 DNA CLEAVAGE (BY SIMILARITY).
 SQ SEQUENCE 759 AA: 85438 MW: 3D862412D72946BD CRC64:

Query Match 32.7%; Score 1711; DB 1; Length 759;
 Best Local Similarity 43.5%; Pred. No. 4,1e-103;
 Matches 364; Conservative 114; Mismatches 261; Indels 98; Gaps 13;

OY 6 VKRVLCVAEKNDAAKGIADLLSGMRNRREGLSKFNKIYEDYHLYGONTVMWTSVSGH 65
 DB 1 MKRALPVAEKNDAAKGVAAALISNGTANRRGRSKFNKIYLTNLELQQAISVTSVSGH 60
 OY 66 LLAHDQMGFRKMGSCNPVLFEAEIEKYCPENFVDIKTLERETGQCALVMTQDRE 125
 DB 61 MMNFOEHENKSNNOJTSAMELFRAPRHVVTPEMKLEIYLRBOAQHDILVWTDQDRE 120
 OY 126 GENIGFEIIVHCVAKVPNLOVLRARSEITPHAVRTACENLTPEPDVSADAVROEIDL 185
 DB 121 GEALIGAEIYKVCDSNRRLDIFRARSEITKAITRAARLILNDETVAADVCRSLDL 180
 OY 186 RIGCAFTTRQTLRLQRIPEVLAE---QLISGSCOPPLGFGVVEREKAIQAFVEIPIH 241
 DB 181 RIGSAFTRLQTLRLRNFRDLGONDTSOVYSYGSCOPPLGFGVTDRYKMIENFVSPEFW 240
 OY 242 RIKVTDHKQGIYEFNMKRRLFNHTACVLVQLQVEDPMATVVEVSKSKMRRPALD 301
 DB 241 KILVETRESHKVEFLMDRRRLFDROTVDILHDECKETKAHAEKVAKKSKMRPALD 300
 OY 302 TVLEKILASRLIRINKETMRIAEKLYTQGISYPTETNIFPRDLNLVLEVGQTPDPR 361
 DB 301 TVLEKILISKLRMSAKQTMQVAKLYSKGISYPTETKPKFAGLNLPLVVOQQTOSNI 360
 OY 362 WGAFAOSILERGGTPRNGKSDQANPPIHPTYT--NNIQGBQRLYEYIVHFLACCS 419
 DB 361 WGFPAFVEILO--NCGVNPNGRKSDEAPPIHPLKTEKHQIQGQDWKYYELVVRFLACVS 419
 OY 420 QDAQGGETVEIDIAOERFAHGLMILARNYLDVYPRDHSKILPYEAGSHFQSTVE 479
 DB 420 QDAQGGETVNLVYGEKFPASGLRIKMGYLVYVYEKGNRLPYTGTGERPTDELK 479
 OY 480 MVDGETSPKLLTEADLIALMERHIGTDTAHEHTETIKARMYVGLTPDKRPLGHILM 539
 DB 480 IGGGKTQAPDFLEADLISLMDKYGTGDTAHEHEIKITREYIGVRPQKILPSPLGL 539
 OY 540 GLVEGTDSCGYEMSKPDLRALEADLKLJDDGKKDFVVLROOYQKQYVIEAVAKAK 599
 DB 540 ALVDGVDMDKGFAMSKPDLIRANLEIGLEICIDGRKOKEVDEQIGKRAIFVESERRIGV 599
 OY 600 LDALAOYFNGTELAQOEEDIVPAMPEPIKPCQCNKDWLTKRKNGGFLSCGGEFECR 659
 DB 600 LSGSLQRYLIDKNNQAGG---PGGP-----GGG-----GGP--- 627
 OY 660 SAVMLPDSVLEASRSSVCPVCPHPYRLKLFKRGSLPPTMLEPFCVCCIGGCD--- 715

DB 628 -----PRGPGGGGGGGTGPAPPPKPPKRRGRPPRKSISPAVK-----NGHDEPND 675
 OY 716 ---TLREILDLRFGSGPPRASQPSGRLOANOSLNRMDNSCHPOPADSRGTGSSKALAOQL 772
 DB 676 TIVLTLSV-----FGSMNPPKAPKPRAPKRSAA----- 703
 OY 773 PPPTAGESNSVNCNGQ---EAVLLTVRKRGPRGRGPFPC-----NGSGCNFLWA 822
 DB 704 -APKEDEEEVEFCQCEPRRAVTKVVKQKPGNKGKRYTCSLPYTSSEKCNFFKWA 759

RESULT 5
 TOP3_SCHPO
 ID TOP3_SCHPO STANDARD; PRT; 622 AA.
 AC 060126;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA topoisomerase III (EC 5.99.1.2).
 GN TOP3 OR SPBC1665.12C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA MEDLINE=20042440; PubMed=10572171;
 RA Mafiah M., Han C.S., Langston L.D., Hope J.C., Zlgours N.,
 RA Freyer G.A.;
 RT "The top3+ gene is essential in Schizosaccharomycetes pombe and the
 RT lethality associated with its loss is caused by Rad12 helicase
 RT activity.";
 RL Nucleic Acids Res. 27:4715-4724(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Beck A.,
 RA Reinhardt R.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
 CC DNA, followed by passage and rejoining.
 CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF126287: AAD2485.2: -
 DR EMBL: AL023554: CAI19038.1: -
 DR InterPro: IPR003601: DNATopI_ATP_bind.
 DR InterPro: IPR003602: DNATopI_DNA_bind.
 DR InterPro: IPR000380: Pro_topoisomerase.
 DR InterPro: IPR002936: Toprim.
 DR Pfam: PF01131: Topoisom. bac. 1.
 DR Pfam: PF01751: Toprim. 1.
 DR PRINTS: PR00417: PRTPISMRASE1.
 DR SMART: SM00437: TOP1AC. 1.
 DR SMART: SM00436: TOP1BC. 1.
 DR SMART: SM00493: TOP1RM. 1.
 DR PROSITE: PS00396: TOPOISOMERASE_I_PROK. 1.
 DR Isomerase: Topoisomerase: DNA-binding.
 FT ACT_SITE 330 330 DNA CLEAVAGE (BY SIMILARITY).
 SQ SEQUENCE 622 AA: 71169 MW: EF291B1717B586C0 CRC64:

Query Match 24.5%; Score 1279; DB 1; Length 622;
 Best Local Similarity 42.2%; Pred. No. 2,6e-75;

	Matches	253: Conservative	127: Mismatches	198: Indels	22: Gaps
QY	8	KYLCAAEKNDAAKAGIADILSNGRNRREGLSKFNKTYEFYHLYGVN--	--TMYTSYSG	64	
Db	2	RYLCAAEKNKSIASKVSLIGGVHVRNRRTKRYKKNYKNDFFSN -GGVNSDDYMTSYSG		60	
QY	65	HLIADPQMQRKQSCAPLVLEAEIEKYCPENFVDIKKLEETRFCCQLVITWTCDR		124	
Db	61	HLTEASFPESEVSSMSVQDVLFRDQIITTSYSKAEVLAADIKKEARNAQVLYITWTCDR		120	
QY	125	BGENIGFEI IHVCKAVKRNLOVLRASEITPRAVFLACENLTERDDORSADVADVROELD		184	
Db	121	BGEHIGVEISNARASNSPSIOVIRADFNNLERSHISAKKRPDVSKNADAVAREIELD		180	
QY	185	LRIAGAFRFOFLRQRIFFPELVLEQLSYSCGCPPTLGEFVERPKALQIAPEIFERIK		244	
Db	181	FRUGAFTRLOTLQLOKSF -DLONKIISTYGPCPFPPLIGFVVDWMQVEDVPEFTWHLR		235	
QY	245	VTHDHKDIGIVEFNMKRRLFNHTACLYVLQCVDPDPAVTVEVSKSRKPMQALDTVE		304	
Db	240	FVDRKQGTGTYQNMERAKVPRDLTMTLIELNCBCKTAKVANNITQKPKTKYPLSTVE		299	
QY	305	LEKLASRLKLRINAKTMTIAEKLYTQGISYSPRETINFFPBDLVLYLVEDQTPDPWGA		364	
Db	300	LTKLPCKHLRISAKTTLLELANLTYNGVSPYKRTEDQFDDSNMLHAIQKLTGAQEMDS		355	
QY	365	FAOSLIEGSGPLPRGNKNSDOAHPIHPTKTKTNL--QGDEQLSYFIYVHAFACSD		421	
Db	360	YAEGLLADYRPRKRGKHNDAHPRIHVQVWHSALRPSQDHMKVVELITRFLACSDN		419	
QY	422	AGOGETTYEIDIAQREFAHGLMLIARYLDVYTYDHN -SPKILPYEGSHFOPSTVEN		480	
Db	420	AKGATTLVQVMEELDFSKGLDLTEKNYLVLYPYEKMESSDQLEPYRLHEEPFLHDM		479	
QY	481	VDGEFSPKLLTEADILALMEKHGIGTATPAEHIEITKARMY-----GLTPDKRF		533	
Db	480	MDSSTSSSYTEPERLILMDANGCIDTATMAHIEKYQEREYLIKRRKKGQVY---EF		538	
QY	533	LPGLHGMGLVEGYDSMGYE--MSKRPDALEADIKLICDGKKDFYVLNQOYQKYQVF		590	
Db	537	VPSSIGVALALAGYDEIGLEMSLTPRFLEKMEVOLKNIENGQLNRNLVHMLITQGFEDVF		596	
RESULF 6					
TOP3_YEAST					
AC	PI3099;	STANDARD;	PRT;	656	AA.
DT	01-JAN-1990 (Rel. 13, Created)				
DT	01-JAN-1990 (Rel. 13, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	DNA topoisomerase III (EC 5.99.1.2).				
GN	TOP3 OR EDRI OR YLR234W OR L6083.3.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
NC	NCBI_TaxID=4932;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89324087; PubMed=2546682;				
RA	Wallis J.W., Chretien G., Brodsky G., Rolfe M., Rothstein R.;				
RA	"A hyper-recombination mutation in S. cerevisiae identifies a novel				
RT	eukaryotic topoisomerase";				
RL	Cell 58:409-419(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=528C / AB972;				
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,				
RA	Favellio A., Fulton L., Gattung S., Greco T., Kirsten J.,				
RA	Kucaba T., Hallsworth K., Hacking J., Hillier L., Jier M.,				
RA	Johnson D., Johnston L., Langston Y., Latreille P., Le T.,				
RA	Mardis E., Meneses S., Miller N., Nhan M., Pauley A., Peluso D.,				
RA	Ritken L., Riles L., Taich A., Trevasakis E., Vignati D.,				
RA	Wilcox L., Wohltman P., Vaudin M., Vaudin R., Waterston R.;				

```

RL submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 578-592; 605-615; 794-805 AND 1038-1046.
RX MEDLINE=93330283; PubMed=8336724;
RA Jiang W., Middleton K., Yoon H.-J., Fouquet C., Carbon J.:
RT "An essential yeast protein, CBF5p, binds in vitro to centromeres and
   microtubules."
RL Mol. Cell. Biol. 13:4884-4893(1993).
CC -I- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER. ESSENTIAL
CC FOR PROPER CHROMOSOME SEGREGATION IN BOTH MEIOSIS AND MITOSIS.
CC WEAKLY RELAXES NEGATIVE SUPERCOILS AND DISPLAYS A DISTINCT
CC PREFERENCE FOR BINDING SINGLE-STRAINED DNA. INTERACTS WITH SGS1.
CC THE TOP3-SGS1 PROTEIN COMPLEX MAY FUNCTION AS A EKAROTIC REVERSE
CC GYRASE INTRODUCING POSITIVE SUPERCOILS INTO EXTRACHROMOSOMAL
CC RIBOSOMAL DNA RINGS.
CC -I- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC DNA, followed by passage and rejoining.
CC -I- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/TII TOPOISOMERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M2A939; AAA35161.1; -.
DR EMBL; U19027; AAB67406.1; -.
DR PIR; A33169; ISBT3.
DR SGD; S0004224; TOP3.
DR InterPro; IPR003601; DNATOP1_ATP_bind.
DR InterPro; IPR003602; DNATOP1_DNA_bind.
DR InterPro; IPR000380; Pro.topoisomerase.
DR InterPro; IPR002936; Toprim.
DR Pfam; PF01131; Topoisom_Bac; 1.
DR Pfam; PF01751; Toprim; 1.
DR PRINTS; PR00417; PRTPISMRASEI.
DR SMART; SM00437; TOPIAC; 1.
DR SMART; SM00436; TOPIBC; 1.
DR SMART; SM00493; TOPRIM; 1.
DR PROSITE; PS00396; TOPOISOMERASE_I_PROK; 1.
KW Isomerase, topoisomerase, DNA-binding.
KM Act_site 356 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 656 AA; 74370 MW; 51DF78936A8B4F4 CRC64;
Query Match 22.5%; Score 1174.5; DB 1; Length 656;
Best Local Similarity 41.8%; Pred. NO.1.6e-68;
Matches 264; Conservative 104; Mismatches 213; Indels 51; Gaps 12.
QY 8 KVLCAEKNDAKGIADLLNSGRMRREGLSKFNKIYEFDYL-----GQNTVMVMSV 62
    ||||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 KVLCAEKNSIAKAASQLIGGGRSTRSDGYVYKNYDPMFGFPFARGANGACEVTMTSV 61
QY 63 SGHLLAHFOQMQRWQSCNPLVEAEILEXCPENVDIKTLRETRPCOALVIIMDC 122
    |||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 AGHLTGIDFSHDHSWGKCALOELDPALNETMNNOKKIASNIREANADYLMITWDC 121
QY 123 DREGENIEFEIHWCK--AAKPNIQVLRAPSETITPAHVRFASCENLTPEPDVSADV 179
    |||| | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 DREGYIOWELMOEQEKRGNRILONDQYRVAFVSHEROHILMAARNPSRLDKSVHYAGT 181
QY 180 ROELDRLIGAFTFRQT-----LALQRIPEPVLA-----QLISGCORPTAGF 224
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 182 RIIEDLRAGVFETRLITETLRNKLENQATMTMKDGAKKHGKNKNDQVVSYGCEPPIEG 241
QY 225 VVERKALQAVPELFHFHKIKTHDKD--GIYEFNMKKRHRLNHRACLYLVLCYEDP-- 280
    || : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 242 VDDEPRFERLNRYVEPFWTYIQLVYEKKDNGGITTFQWDKGHLFDRLSYLTFYTCLETAGN 301
QY 281 MATVEVNSKRSKRRPOLDTVELEKLASKRLINAKETMRIAEKLYTGQIYSYPTET 340

```

```

DR 302 VAGVVDLKSPTTKYRPLPTTVELQKNCAHYRLNANKOSLDIAEKLYQKGFISYPTRET 361
OY 341 NTFPRDLNLTVEQOTP-----DPRMGAFASQILTEGCTP-----PRNGKSKDQ 385
DB 362 DTFPHAMDLSKLVKQAOQLDOLAAGRTAWASVAASTLQPEPNTSNNKFKFPRSGSHDDK 421
OY 386 AHPPIHPTKYTN---NLQGDQRLEFIVRHFLACCSODAQOGQETVEIDIAOERFAHG 442
DB 422 AHPPIHPTKYTN---NLQGDQRLEFIVRHFLACCSODAQOGQETVEIDIAOERFAHG 442
OY 443 LMLANRYLDVYPPDHV-SDKILPVYEOGSHFQPSVTEWDEGSETPKLTLEADLALME 501
DB 482 LVYLENPLFDVYPMARWETTKOLPRLEMANLVDAIAKEMKAGTTPAPKPMTESELILMD 541
OY 502 KKGIGDATHAEHIEFTIKARMY---GLTPDKFLECHLGMGLVEGYSKGYE--MSKPD 556
DB 542 TNGIGDATHAEHIDKIOVNVYSEKVEKTYLQPTTLGSLVHGFEALGLEDSFAKPF 601
OY 557 LRAELADLKLICDGKKDFVVLROQYOKYKQ 588
DB 602 QRREMEDLKKICEGHASKTDVVKDIVEKYRK 633

RESULT 7
TP3B_MOUSE STANDARD; PRT; 862 AA.
ID TP3B_MOUSE STANDARD; PRT; 862 AA.
AC Q92321;
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DE DNA topoisomerase III beta-1 (EC 5.99.1.2).
GN TOP3B1 OR TOP3B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-Testis;
RX MEDLINE=9003190; PubMed=9786843;
RA Seki T., Seki M., Onodera R., Kanada T., Enomoto T.;
RT Cloning of cDNA encoding a novel mouse DNA topoisomerase III (Topo
RT IIIbeta) possessing negatively supercoiled DNA relaxing activity,
RT whose message is highly expressed in the testis."
RL J. Biol. Chem. 273:28533-28556(1998).
CC -1- FUNCTION: POSSESSES NEGATIVELY SUPERCOILED DNA RELAXING ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC DNA, followed by passage and rejoining.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS.
CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AB013603; BAA34227.1; -.
DR HSSP: P06612; 1ECL.
DR MGD: MGI:133803; Top3b.
DR InterPro: IPR003601; DNATopI_ATP_bind.
DR InterPro: IPR003602; DNATopI_ATP_bind.
DR InterPro: IPR000380; Pro_topoisomerase.
DR InterPro: IPR002936; Top1im.
DR Pfam: PF01131; Topoisom_Dac; 1.
DR Pfam: PF01751; Top1im; 1.
DR PRINTS: PR00417; PRTIPSMRASET.
DR SMART: SM00437; TOP1AC; 1.
DR SMART: SM00436; TOP1BC; 1.

```

```

DR SMART: SM00493; TOP1RM; 1.
DR PROSITE: PS00396; TOPOISOMERASE_I_PROK. 1.
KW Isomerase; Topoisomerase; DNA-binding.
FT ACT_SITE 336 336 DNA_CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 862 AA; 96949 MW; D2C05429F79FD5CC CRC64;
Query Match 20.6%; Score 1078; DB 1; Length 862;
Best Local Similarity 36.0%; Pred. No. 3.9e-62;
Matches 276; Conservative 112; Mismatches 311; Indels 68; Gaps 17;
OY 6 VRRVLCVAERNDAAKGIADLLSNGRMRRBGLSKFNKTYEFDYHLXGONTMTVTSYGH 65
DB 1 MKTYLVAAEKPSLAQSTAKLISRGNMSSHGLNACSVHRYTGTFAGQVPHFKMTSCGH 60
OY 66 LLAHDFOMORRKNWOSCPVLYF-EAELEKTCPEPNVDIKTLERETKQCALVYWTCCDR 124
DB 61 VMTLDFGKYNKMDKVPDAELFESQAPPEKREANRLNMFVLQVEGCGCDVYVLMDCDK 120
OY 125 EGENIGFEIIVHCKAVPNLQ-----VLRARFSEITPHAVRTACENTLEPDQRYSDAVD 179
DB 121 EGENICEVLDVAVLPVNNMANGKTYFRARFSSITDTIDCNAMTRLSEPHNEALSVD 180
OY 180 ROELDLRIGAFTRFQTLRLQRTFPEVLADQLISYSCQEPFLLGVVEREKAQAFYBEI 239
DB 181 ROELDLRIGAFTRFQTLRLQRTFPEVLADQLISYSCQEPFLLGVVEREKAQAFYBEI 239
OY 240 FHRIRV-TTHDKDGIIVEFNKRHRLEFNHTACLVLYQLCEDPMATVVEVNSKPRSKRPQ 298
DB 240 YWVLQARVHTDKDEESLLDDMDRVRFDMETLQMLNTKLEKAMVATSKRKAORPL 299
OY 299 ALDVELEKLSRLRLINAKETMRLIAKLYTQGIYSYPRTEINFPDNLNLTVEQOTP 358
DB 300 ALNTEMLRVAASSALGMPQIAQIAERLYTQGIYSYPRTEINFPDNLNLTVEQOTP 359
OY 359 DPRMGAFASQILTEGCTPPRNGKSKDQAHPIHPTKYTN---NLQGDQRLEFIVRHFLA 416
DB 360 HPYVADSVKOLLACGINRPRKGHGAGD-HPIITPMKSTAELEGDAMRLLEYITRHHIA 418
OY 417 CCSODAQOETVEIDIAOERFVAHGLMILARNLYDVPYDHS----DKILPVYEOGSH 472
DB 419 TVSHDCKYLLSTISFRIGPEHFTCMGTVISPGTEIMP--MQSVPLEESLPTCOKGDT 475
OY 473 FQPSVTEWDEGSETPKLTLEADLALMEKHGIGTDTHAEHIEFTIKARMYVGLTPDKRF 532
DB 476 FTVGEVVMLEKQISPPRYLLEALITLMEKHGIGTDTHAEHIEFTIKARMYVGLTPDKRF 535
OY 533 LPGLHGLMGLVEGYSKGYECPDRLRAELADLKLICDGKKDFVVLROQYOKYQVFE 592
DB 536 KPTNLGIVLVHGYKKIDAEVLPTIRSAVEKQMLINQKADYHQVGLTLDIFKRKFHY 595
OY 593 AVAKAKKLDEALAOYFGNGTELAQOEDITYPAMPPIRKCPOCKNDVYLTKKNGGFYLS 652
DB 596 FVVSIAQMDLEMEVSF---SPLA-----ATGKPLSRCCGCHFRMRYQAKPSRLKCS- 644
OY 653 MGPEECSSAWMLPDS-----VLRASDSS--VCPVQDPAPVYRLK 690
DB 645 ---HCDEYTYTLTPNGTIKYLKELRCLPDLDEFLVWSSGSKGYPLCPYCYNHPPR-D 699
OY 691 LKFKRGLSPTMPLEFYCC-----IGGCDDTLREILDRFSGGP 729
DB 700 MKKGMGCGNECHP---TCQHSLSMLGICGCEBENGVLVLDPTSGP 742

RESULT 8
TP3B_HUMAN STANDARD; PRT; 862 AA.
ID TP3B_HUMAN STANDARD; PRT; 862 AA.
AC Q95985; Q95985;
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 01-MAR-2002 (Rel. 41, last annotation update)
DE DNA topoisomerase III beta-1 (EC 5.99.1.2).
GN TOP3B1 OR TOP3B.
OS Homo sapiens (human).

```


CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99128286; PubMed=9927731;
RA	Ng S.-W., Liu Y., Hasselblatt K.T., Mok S.C., Berkowitz R.S.;
RT	"A new human topoisomerase III that interacts with Sgs1 protein.";
RL	Nucleic Acids Res. 27:993-1000(1999).
RP	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=91228902; PubMed=9074928;
RA	Kawasaki K., Minoshima S., Nakato E., Shibuya K., Shintani A.,
RT	Schmeles J.L., Wang J., Shimizu N.;
RL	"One-megabase sequence analysis of the human immunoglobulin lambda gene locus".
RP	Genome Res. 7:250-261(1997).
RP	[3]
RP	SEQUENCE FROM N.A.
RA	Riou J.F., Goulaouic H., Grondard L.;
RL	Submitted (Feb-1999) to the EMBL/genbank/DBJ databases.
RP	[4]
RP	SEQUENCE FROM N.A.
RC	TISSUE-Skin;
RA	Strausberg R.;
RL	Submitted (Feb-2001) to the EMBL/genbank/DBJ databases.
CC	- I - FUNCTION: POSSES NEGATIVELY SUPERCOILED DNA RELAXING ACTIVITY.
CC	- I - CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded DNA, followed by passage and rejoining.
CC	- I - ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY HAVE DIFFERENT C-TERMINAL DOMAINS.
CC	- I - TISSUE SPECIFICITY: THE TOP3B ISOFORMS HAVE DIFFERENT TISSUE SPECIFICITIES. ISOFORM 1 IS FOUND IN TESTIS, HEART AND SKELETAL MUSCLE. A 4 KB TRANSCRIPT, WHICH PROBABLY REPRESENTS ISOFORM 2, IS FOUND IN THYMUS, KIDNEY AND PANCREAS.
CC	- I - SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sdb.ch/announce/ or send an email to license@sdb-stb.ch).
CC	-----
DR	EMBL; AF053082; AAD15791.1; -
DR	EMBL; AF017146; AAD01614.1; -
DR	EMBL; AF125216; AAD29670.1; -
DR	EMBL; BC002432; AAH02432.1; -
DR	HSSP; P06612; IECL.
DR	MIM; 603582; -
DR	InterPro; IPRO03601; DNATOP1_ATP_bind.
DR	InterPro; IPRO03602; DNATOP1_DNA_bind.
DR	InterPro; IPRO00380; Pro_topoisomerase.
DR	InterPro; IPRO02936; Toprim.
DR	Pfam; PF01131; Topoisom_bac; 1.
DR	Pfam; PF01751; Toprim; 1.
DR	PRINTS; PR00417; PRTPISMRSET.
DR	SMART; SM00437; TOP1AC; 1.
DR	SMART; SM00436; TOP1BC; 1.
DR	SMART; SM00493; TOPRIM; 1.
DR	PROSITE; PS00396; TOPOISOMERASE_1PROK; 1.
DR	Isomertase; Topoisomerase; DNA-binding; Alternative splicing.
FT	ACT_SITE 336 336
FT	VARSPLIT 703 730
FT	DNA_CLEAVAGE (By SIMILARITY). GMSGNECHPSCSLSLMGLGGQVEEC -> GCSCSHLS TGSCSLSPVPALHOAGL (IN ISOFORM 2).
FT	MISSING (IN ISOFORM 2).
FT	GMSGN -> VVPCV (IN ISOFORM 3).
FT	MISSING (IN ISOFORM 3).
FT	D -> N (IN REF. 4).
FO	SEQUENCE 862 AA; 96661 MW; 75352827856CFEPR CRC64;

```

Query Match Similarity      20.4%: Score 1066; DB 1: Length 862;
Best Local Similarity      36.3%: Pred. No. 2, 3e-61;
Matches 280; Conservative 109; Mismatches 306; Indels 76; Gaps 19;

QY      6 VRKVLCAEAKNDAAKGIADLLSNGMRREREGSKFKFIYEFYHLYGQAVTMYMTSVSGH 65
      1 KMTVLVMAKPKPSLAQSIAMKLSRGLSSHKGILNGACSVNHEYTGFAGQVYRRKMSVCGH 60
QY      66 LLANDFOMQFQWOSCNPLVLR-EAEITEKYCBENFYDIKKTLETRTQCOALVYITDCDR 124
      61 VMTLDFELGYNMKWDKDPALPELFSQAETKEKANPKLNMVKKFLQVEBGRGDYIVLMDCKK 120
QY      125 EGENIGETLIHHCKAQNPMYIQ-----VLARFSEIRPHAVRTACENTLEPDDQVSDA 176
      121 EGENICEFVLE---DAVLPRYMKAHGGEKIVFAFRSSITIDTICNMAAGLERDHNELAS 177
QY      177 VDVRQELDLRIGAAFTREFQTLRLQRIFFEVLAEBOLISYSGSCOPPLGFEVERFKAIQAFV 236
      178 VDARQELDLRIGCAFTREFQTKYFGQKYG-DLSSILISFGCPQTLPGFCVERHNDKIQSPK 236
QY      237 PELFH-RKLYVNHDKDGIYVERNMKRNHLENFTACLYIQLCEVDRPMATVYVERSKPSK 294
      237 PELYVWLQAKVNTD-KDRSLLDMDQVRVRFDEBIAOMFLNMKLEKAEAVETSREKRAK 295
QY      295 MRPQALDYVELLEKLSARKIRINAKETMRFAEKLTYOGYISYPTRETNIPRLNLTAVE 354
      296 QRPRLNLTVEMLRVASSSLGCMGRQAMQVAEBLYOGYISYPTRETYHPEMFDLKGSIR 355
QY      355 QQTPPDRMGAFQASQILIERGGCPTRPNQNSDQAHPRIHPTKYTN--NLQGEORLYEFTVR 412
      356 QQANHPYVADTVKRLLAEGINRRPKGHADG-DHPRTPKKSAATEALGGDAMRLYEYIYR 414
QY      413 HPLACSOAOAQOQPTVEIDIAQERVANGMLILARNYLDVYRUDHWS---DKLLPYVE 468
      415 HTIATVYSHDCKLYQSTIISRTIRPELFTSGKTVLSGFEYVMP---WQSVPLEESLPTCQ 471
QY      469 QGSHPQPTVEVWDETSPKLLTEADLLALAMEKHGIGTADATNAHEITIKARMYVGLTP 528
      472 RQDAFQVGEVKNLEKQTNPDYLTAEELITLMEKHGIGTADASIPVHINNIGQRYVYVES 531
Db      529 DKRFLPGHLMGVLGEGYDSMGYEMSKPDLRAELADKLICGKKDKFVYLROOYQATKQ 588
      532 GRRLPTNGLIVLAVGYKKIDAEILPPTIRSAVERQQLMLIAQGKADRYVLEHNTLDVEKR 591
QY      589 VITEAVAAKKKIDELADLAQYFGNGTELAQOEDITYPAMPPEIRKCPQCNMDYLTKTKNGGF 648
      592 KHHYVVDISLACDELMEVSEF---SPLA-----ATGPKLSCGCKCHRPMKYIAQKPSRL 641
QY      649 YLSCMGFPECRSAAVWLPDS-----VL--EASRDS--VCPYVCOAPHP 686
      642 HCS-----HCDETTYLLPQGTIKIKYELKRCPLDDELYLWSSGSGKSYPLCPYCYNHP 696
Db      697 FR-DMKKKGWGCNECTHP---SCQHSLSMLGIGQVCESEGYVLADPTSGP 742

RESULT 9
TP3B_DROME STANDARD: PRT: 875 AA.
AC 096651; Q9W416;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA topoisomerase III beta (EC 5.99.1.2).
OS TOP3-BETA OR TOP3 OR CG3458.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

```



```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (Relaxing enzyme)
DE (Unwinding enzyme) (Swivelase).
GN TOPA OR TA0063.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
OC Thermoplasma
NCBI_TaxID=2303;
OX NCBI_TaxID=2303;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Granel W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Luppas A.N., Baumeister W.;
RT "The genome sequence of the thermophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513(2000).
CC -1- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC DNA, followed by passage and rejoining.
CC -1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AL445063; CAC11211.1; -
DR InterPro: IPR003601; DNACoPI_ATP_bind.
DR InterPro: IPR003602; DNACoPI_DNA_bind.
DR InterPro: IPR003603; Pro_topoisomerase.
DR InterPro: IPR002936; Toprim.
DR Pfam: PF01131; Topoisom_bac; 1.
DR Pfam: PF01751; Toprim; 1.
DR Pfam: PF01396; zi-C4_topoisom; 1.
DR PRINTS: PR00417; PRTPISMRASE1.
DR SMART: SM00437; TOP1AC; 1.
DR SMART: SM00436; TOP1BC; 1.
DR SMART: SM00493; TOPRIM; 1.
DR PROSITE: PS00396; TOPOISOMERASE_I_PROK; FALSE_NEG.
KW Isomerase; Topoisomerase; DNA-binding; Zinc-finger; Metal-binding;
KW Repeat; Complete proteome.
FT ZN_FING 611 638 C4-TYPE 1.
FT ZN_FING 673 700 C4-TYPE 2.
FT ZN_FING 719 744 C4-TYPE 3.
FT ACT_SITE 312 312 DNA_CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 770 AA; 87667 MW; 75DA8DD7BC3B8A22 CRC64;

Query Match 12.5%; Score 653; DB 1; Length 770;
Best Local Similarity 26.5%; Pred. No. 1e-34;
Matches 205; Conservative 130; Mismatches 329; Indels 110; Gaps 22;

```

```

Db 113 EGGLEVEALDIIKEGEE--EIRAKFSALTKNKNEILDSFKNLIGVNSLADADARESID 170
Qy 185 LRIGAFTFRPOTLRLORIPEVLAEOILSYSCQPTPLGVEERFKAQAFVPEIEHRK 244
Db 171 LINGSVLTFRFVSWTGR-----LGKSFLSAGRVQTPTLAIVNDREREIOQFPERERYWTIS 225
Qy 245 VTHDHKQGYIEFPMKRRRLRFLNHTACVLVXOLGVEDPAAVVEVRSKSKMRQALDVE 304
Db 226 ITFD-KDQGRKARIPEIKQDDTA-EKIVE-ALKGKGRVSYTSKSDHDIRRPAPESTTE 282
Qy 305 LEKLASRKLRIINAKETRIAEKLYTQGISYPTETNIFPPDNLTVLVEQGTAPPDPMGA 364
Db 283 FLREASR-ICIMPTKAMSIENLYMRGLISYPRDNTVYRSINLKSVL-KLEENTAYSK 340
Qy 365 PAOSILERGGPTPRNGKSDQAHPIHT-KYNNLOGDEQRLEYFVIRHPLACCSDDAQ 423
Db 341 YKKEIEFDFRILFSRGRIFETTDHPPIVPDPSPKQLGDRVYDLILRHFLSTLYRDGK 400
Qy 424 GQETTVETIDIAQERFVNHGMLILARNLDVYPVDMNSDKILPYEGOSHQPSTVEAVDG 483
Db 401 KYVAEALIVNGYTFKAGQHTTDRGWTETIGYDP-KDVIPLTEBEDLKAIDMNTQRE 459
Qy 484 ETSPPKLLTEADLIALMEKHGIGTDATNAEHIETIKARMYGLTPDRRFLPGHLMGLVE 543
Db 460 ETRPPRYDMSSLLKKMEELNLGKSTRHDIGLIERGFEGNPKV---PTPLGMAFID 516
Qy 544 GYDSMGYEMSKPDILRAELADLKLICDQKKDFVLRQOVYKRYOPIEAVAKAKKIDEA 603
Db 517 AVRSVNSHIADPEWTKARLEEDMDRIEKNEMSKNDVSESKMLHEVLSHFLTAKYADVKI 576
Qy 604 LAQYFGSGTEL-----AQGEIYVAMPPIR-----K 630
Db 577 ITKGINAGQELIGDCPFHEGKKNIVIRDRFTYVACED--PSCIKNFIRKNGSITLSDQK 634
Qy 631 CPQC-----NKD-----AYLTKRKNNGGYLSC 652
Db 635 CPVCGELPMIKIRKQSPETKICIDPDSCSYNRENEDYCEPADHGRVLVLRQSKRFLGCG 694
Qy 653 MGPEECRSANVLPDSVLEASRDSSVCVCPHPYRKLKFKRSLPTPLPLEF 706
Db 695 SNYPKCTVYPLP-QMGRIRTKTGEVCPGA-PILALSRNGKMKFCFPMNOCEY 746

RESULT 12
TOP1_PYRAB STANDARD; PRT; 685 AA.
ID TOP1_PYRAB STANDARD; PRT; 685 AA.
AC Q9UY58;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (relaxing enzyme)
DE (unwinding enzyme) (Swivelase).
GN TOPA OR PAB1430.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Helling R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT structure and evolution."
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC DNA, followed by passage and rejoining.
CC -1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE

```

```

CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AJ248287; CAB50334.1; -.
DR      HSSP: P06612; 1ECL.
DR      InterPro: IPR003601; DNATopI_ATP_bind.
DR      InterPro: IPR003602; DNATopI_DNA_bind.
DR      InterPro: IPR000380; Pro_topoisomerase.
DR      InterPro: IPR002936; Toprim.
DR      Pfam: PF01131; Topoisom_bac; 1.
DR      Pfam: PF01751; Toprim; 1.
DR      Pfam: PF01396; Zf-C4_Topoisom; 1.
DR      PRINTS: PR00417; PRTPISMRASEI.
DR      SMART: SM00437; TOP1AC; 1.
DR      SMART: SM00436; TOP1BC; 1.
DR      SMART: SM00493; TOPRIM; 1.
DR      PROSITE: PS00396; TOPOISOMERASE_1_PROK; FALSE_NEG.
KW      Isomerase; Topoisomerase; DNA-binding; Zinc-finger; Metal-binding;
KW      Repeat; Complete proteome.
FT      ZN_FING      611      C4-TYPE 1.
FT      ZN_FING      611      C4-TYPE 2 (ATYPICAL).
FT      ZN_FING      658      682      DNA_CLEAVAGE (BY SIMILARITY).
FT      ACT_SITE      317      317
SQ      SEQUENCE      685 AA; 78195 MW; 4888C43AB87B485D CRC64;

Query Match      11.5%; Score 600.5; DB 1; Length 685;
Best Local Similarity 27.1%; Pred. No. 2.1e-31;
Matches 200; Conservative 129; Mismatches 297; Indels 111; Gaps 22;

QY      9 VLCAEKNDAAKGADLDSNGRMRRREGLSKFNKIYEDVLYQGNVTMTVSGHLLA 68
DB      2 ILVIAEKPNARKTAGALSERPRTRK--TIFGVY--YEFEDGKKLYIAVGHLYG 55
QY      69 HDFOFOM-----RKQSCNPLVFEAEIEKYCPENFVDIKTLERETROCAQALVTMT 121
DB      56 LAPQDFGVPIPIDEM-----VPVYIAEKMKDYAKDYI---KALSLVAKKVEIYACD 107
QY      122 CDREGENIGEITIHVCAVAPNLOVLARSETITPHAVTACENLTER--DQVSDAVDV 179
DB      108 YDEGEVIGYALTKYACGVPS--RAKRMKESALTFRDLNAMRNL--EPTINFGMANAGIA 165
QY      180 ROELDLRIGAAFTFRPQLRLQRIPEVLAEQLSYGSCQPTLGFFVEREFAIAFVPEI 239
DB      166 RHVIDMWGNLSALHAIKRASGKNV--VLTGKRYQGTFLKFLVEREREIOSFVPKP 222
QY      240 FHRITVTHDKDGIIVERNMKRHRHLNHTACLVLYQLCEDPMATVVEVRSKPSKMR-- 296
DB      223 YWVIAKLVIEKNKGQKLVANYEKDITMSE-----EAKKILVTEVK--KSKARVSN 268
QY      297 -----PQALDYELKELASRKRLINAKETMRIAEKLYTGYISYPTETNINPRD 346
DB      269 IEVKRQRNRPVPPDLGTLTQREASAFGFSFKTLIDNOSLYEKGPTPESQOKLPKN 328
QY      347 LNLTVLEQGTTPDRPMGAPFQOSTLERGGPMPRNKNSQDAHPRIHP--KYTNNLQGE 403
DB      329 LNFLITQNTSRMFQYRPAHALHGLMPELKFVEGKKEDEPAHAIYTPGELIPGQDLSKDE 388
QY      404 QRLYEFTVRFELACSDADQOETTVEIDIAOERFVAHGMILARNLADV--PYDHMSDK 462
DB      389 EKLYDMIVRRFLALFMPDPAVRESVKVITLAGPHKFMLSGARTVAGQMSLVGCKIKRDEV 448
QY      463 ILPVYEOGSHPOSTVMDGETSPPKLLLEADLIAMEKHGISTDAFNAHETITAKRM 522
DB      449 TLPEFFIGEFVRVQLQVIREKKTKTPARYSPAIVIKKMEIDIGTAKTAQILLETLYQSG 508
QY      523 YVGITPKRRLPGHIGLGVGYDSMGYEMSKPDLRAELKLEADLKLIDGKKDKRVVLRLQ 582

```

```

DB      509 YIEGKKSITKVP--LGAKVITLEKYVEPIISVELTRFEREKMDLINOGLTK----- 559
QY      583 VOKYQVETIEAVAKAK--LDBALQYFGNGTLEAQ-----QEDITYPAMPEPIRKPOCN 635
DB      560 ----EEVIEAKVKTKTLLEEKKKRELEIGLEIAIVGVDKDEIKERKPIYVGCPRCG 615
QY      636 KDMVLK--TKKNGGYLSCMGFPEECRSAAVLPDVSLEASRDSVCPYQCPHPHYRLKLKFK 694
DB      616 GDLIVKYNKKKTKGKRFVGCSSNMPKC-----DVYTPILD----- 647
QY      695 RGSLEPTMPLFEVCCIG 711
DB      648 KGEIIP--NKTCONG 661

RESULT 13
TOP1_PYRHO
ID TOP1_PYRHO STANDARD; PRT; 686 AA.
AC O58336;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA topoisomerase I (EC 5.99.1.2) (omega-protein) (Relaxing enzyme)
DE (unwinding enzyme) (Switzerland).
GN TOPA OR PH0622.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; Pubmed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC -!- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC DNA, followed by passage and rejoining.
CC -!- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
CC -!- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/II TOPOISOMERASE
CC FAMILY.
CC -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AP000003; BAA29711.1; -.
DR      HSSP: P06612; 1ECL.
DR      InterPro: IPR003601; DNATopI_ATP_bind.
DR      InterPro: IPR003602; DNATopI_DNA_bind.
DR      InterPro: IPR000380; Pro_topoisomerase.
DR      InterPro: IPR002936; Toprim.
DR      Pfam: PF01131; Topoisom_bac; 1.
DR      Pfam: PF01751; Toprim; 1.
DR      Pfam: PF01396; Zf-C4_Topoisom; 1.
DR      PRINTS: PR00417; PRTPISMRASEI.
DR      SMART: SM00437; TOP1AC; 1.
DR      SMART: SM00436; TOP1BC; 1.

```


This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2002, 13:48:45 : Search time 6334.08 Seconds
(without alignments) 9822.203 Million cell updates/sec

Title: US-09-744-125-1

Perfect score: 2973
Sequence: 1 ggaaccgagcgcgcagatggc.....tagagaagcttggagtcgac 2973

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
------------	-------	-------------	-----------	----	-------------

1	2949.6	99.2	3755	9	HSU43431	U43431 Human DNA t
2	2233.2	75.1	3741	10	AB006074	AB006074 Mus muscu
3	594.6	23.1	132486	9	HSF10178	AL035367 Homo sapl
4	594.6	20.0	3891	3	AF255733	AF255733 Drosophill
5	446	15.0	88718	2	AC097545	AC097545 Rattus no
6	434.8	14.6	203117	10	AL596215	AL596215 Mouse DNA
7	434.8	14.6	213913	2	AC069534	AC069534 Mus muscu
8	433.4	14.6	87835	3	AC005428	AC005428 Drosophill
9	433.4	14.6	157766	3	AC099017	AC099017 Drosophill
10	433.4	14.6	299449	3	AE003663	AE003663 Drosophill
11	429.4	14.4	73907	2	AC018257	AC018257 Drosophill
12	422	14.2	2422	3	AF057032	AF057032 Caenorhab
13	270.6	9.1	2784	9	AF017146	AF017146 Homo sapl
14	270.6	9.1	2793	9	AF053082	AF053082 Homo sapl
15	270.6	9.1	3133	9	BC002432	BC002432 Homo sapl
16	269	9.0	2680	9	AF125216	AF125216 Homo sapl
17	262	8.8	218329	3	CNS07ECB	AL590444 Chromosom
18	257.8	8.7	1869	8	AF126287	AF126287 Schizosac
19	253.2	8.5	110000	2	LMFICHR36_26	Continuation (27 o
20	240.6	8.1	2979	3	AF099909	AF099909 Drosophill
21	238.4	8.0	3267	3	AF061625	AF061625 Drosophill
22	233.6	7.9	2780	10	AB013603	AB013603 Mus muscu
23	195.2	6.6	3543	10	AB045324	AB045324 Mus muscu
24	194.8	6.6	2390	6	AX073290	AX073290 Sequence
25	177.4	6.0	37619	2	AC091210	AC091210 Drosophill
26	172	5.8	245844	2	AC021094	AC021094 Homo sapl
27	147	4.9	169364	2	CEY56A3A	AL132860 Caenorhab
28	136.2	4.6	2435	8	YSCTOP3	M24539 Yeast (S.ce
29	136.2	4.6	2471	6	AX073076	AX073076 Sequence
30	136.2	4.6	41454	8	YSC18083	U19027 Saccharomyc
31	133	4.5	71516	2	AC014220	AC014220 Drosophill
32	133	4.5	305018	3	AE003437	AE003437 Drosophill
33	115	3.9	207184	2	AC020664	AC020664 Homo sapl
34	105.8	3.6	37619	2	AC091210	AC091210 Drosophill
35	95	3.2	10373	2	AC023710	AC023710 Drosophill
36	94.8	3.2	13114	1	AE009905	AE009905 Pyrobacul
37	89.8	3.0	37852	8	SPBC1665	AL023554 S.pombe c
38	77.6	2.6	251700	1	AP000062	AP000062 Aeropyrum
39	77	2.6	11015	6	AX280003	AX280003 Sequence
40	77	2.6	11015	6	AX281190	AX281190 Sequence
41	76.8	2.6	11015	6	AX356459	AX356459 Sequence
42	74.8	2.5	344000	1	TAC1D1	AL445063 Rhinopla
43	74.8	2.5	165027	3	CEY48C3A	AL117203 Caenorhab
44	74	2.5	88718	2	AC097545	AC097545 Rattus no
45	70.2	2.4	958	33	AC085945	AC085945 Giardia i

ALIGNMENTS

RESULT 1
LOCUS HSU43431 3755 bp mRNA linear PRI 12-JUL-1996
DEFINITION Human DNA topoisomerase III mRNA, complete cds.
ACCESSION U43431
VERSION U43431.1 GI:1292911
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Hanai,R., Caron,P.R. and Wang,J.C.
TITLE Human TOP3: a single-copy gene encoding DNA topoisomerase III
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 93 (8), 3653-3657 (1996)
MEDLINE 96195027
REFERENCE 2 (bases 1 to 3755)
AUTHORS Hanai,R. and Wang,J.C.
TITLE Direct Submision
JOURNAL Submitted (15-DEC-1995) Ryo Hanai, Molecular and Cellular Biology,
Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA
LOCATION/Qualifiers
1. .3755

gene	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="17"	
	/map="17p11.2-12"	
	/cell_line="Jurkat T-cell line, ATCC E6-1"	
	178..3741	
CDS	/gene="TOP3"	
	178..3183	
	/gene="TOP3"	
	/note="putative; uses an alternative initiation codon"	
	/codon_start=-1	
	/product="DNA topoisomerase III"	
	/protein_id="AAB03694.1"	
	/db_xref="GI:1292912"	
	/translation="MIFPVAAYALRWLRREDRAFSHAMEMALRGVRYLQVAAEKND AAKGIADLNSGRMRREGSLSEFNK IYEFPYHLVGOAVTWMVSGLHAPHOMF RKMOGCPVLFEAEITEKVCENPVVDIKLEPTEKRCQALVITWDCDRENGEPI IHYKAVKPNLQYLARFSEITPHAVTAENLTERPQRSDAVDNGEIDLRIAGAF TRFOTLQRIPEVLAEQLITSGSCFPPLGVEVERKAIOAVEPEIFRIKVTYHDH KDGIVERNMRKRLFNHTACLVLYQLCEVPMAVYVERKSPKSMRQALDPRVLEBK LASRKLINAKETMARIAEKLVTQGYISYPTETNIPEFDNLVLEEQTPDMVGA AOSILBERGPTPRNGKSDOAHPIHPTKYTNNGDEORLVEFYVHFLACSSODAG GOETVEIDIAOERFVAHGLMILARNLVDVYPDMSGDKILPYEGSHOPSVEVY DGFTSPKILTEADLIAMKHGIGTDAEHIETIKAMVYGTPTDKRFLGHLEHA GLVEGDSMGTENSKPDRLAEADLKLICDGKDKFVVLROOVQTKQYFIEAVAA KRLDEALVOYFGMTDLAQOEDYPPAMPEDIRKCPQCKDMVILKTKNGEFTLSCMF PECRSAVMLPDSVLEASRDSVCPVPCPHEPVYRLKLFKKGSLPTPTLEFVCCIGCF DPTLRITLDRFSGPPRASOPSGRLAANGSLNRMDNSHOPADSHOTSALAOT LPPPTAAGESNYTCNGOGEAVLIVYRKREKGNRROFPCNGGSCNCFELMADSPNPA GGPRLAYRPLGASLCCPBGRTGHIGFGFGPDSGSGTCLCSQPSVTVTVOKDGN KGRHPTCAKPRDQCGCFQWVDENTRATGTSAGAPSWTGDNRKTLSEAKSRPRASS DMGSTAKPRKSCLC HQHPFCPNR"	
	253..3183	
CDS	/gene="TOP3"	
	/note="type I DNA topoisomerase, belonging to E. coli topoisomerase I/yeast topoisomerase III subfamily"	
	/codon_start=-1	
	/product="DNA topoisomerase III"	
	/protein_id="AAB03695.1"	
	/db_xref="GI:1292913"	
	/translation="MEMALRGVRRKVLCAENKDAKGIADLLNSGRMRREGSLSEFNK IYEDYHLVGOAVTWMVSGLHAPDMOFRKMOGCPVLFEAEITEKVCENPV DIKTLERETROQALVITWDCDRENGEPIIHYKAVKPNLQYLARFSEITPHAV VRTACENLTERPQRSDAVDNGEIDLRIAGAFTRFOTLQRIPEVLAEQLITSGS COPPLGVEVERKAIOAVEPEIFRIKYNHDKDGLVEENMRKRLFNHTACLVLYQ LCVEDPMATVEVERKSPKSMRQALDPRVLEBKLEKLTINAKETMARIAEKLVTQ ISYPTETNIPEFDNLVLEEQTPDMVGAPOSILBERGPTPRNGKSDOAHPIH PTKYTNNGDEORLVEFYVHFLACSSODAGOEFTVEIDIAOERFVAHGLMILAR NYLDVYPTDMSDKILPYEGSHOPSVEVYDGFTSPKILTEADLIAMKHGIG TDAEHIETIKAMVYGTPTDKRFLGHLEADLKLICDGKDKFVVLROOVQTKQYF LKLICDGKDKFVVLROOVQTKQYFIEAVAAKRLDEALVOYFGMTDLAQOEDYPP AMPEDIRKCPQCKDMVILKTKNGEFTLSCMFPECRSAVMLPDSVLEASRDSVCP COHPVYRLKLFKKGSLPTPTLEFVCCIGCFDPTLRITLDRFSGPPRASOPSG RLAANGSLNRMDNSHOPADSHOTSALAOTLPPPTAAGESNYTCNGOGEAVLIV YRKREKGNRROFPCNGGSCNCFELMADSPNPAAGPRLAYRPLGASLCCPBGRT GHIGFGFGPDSGSGTCLCSQPSVTVTVOKDGNKGRHPTCAKPRDQCGCFQWVD ENTRATGTSAGAPSWTGDNRKTLSEAKSRPRASSDMGSTAKPRKSCLC HQHPFC PNR"	
	1964	
variation	/gene="TOP3"	
	/replace="a"	
	3354	
variation	/gene="TOP3"	
	/replace="c"	
	3736..3741	
polysignal	/gene="TOP3"	
	896 a 1020 c 1061 g 778 t	
BASE COUNT	896 a 1020 c 1061 g 778 t	
ORIGIN		
Query Match	99.2%; Score 2949.6; DB 9; Length 3755;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 2952; Conservative	0; Mismatches 4; Indels 0; Gaps 0;	

Qy	5	CCGAGCTCAGATGCGCTCCGAGGCGTGCAGAAAGTCTGTGTGGCCGAAAAAAG	64
Db	248	CCGCGCATGATGCGCTCCGAGGCGTGCAGAAAGTCTGTGTGGCCGAAAAAAG	307
Qy	65	ACGGCGCAAGGGATGCGCCGACCTGCTCAACCGTGCATGAGGGGAGAGAGGAC	124
Db	308	ACGGCGCAAGGGATGCGCCGACCTGCTCAACCGTGCATGAGGGGAGAGAGGAC	367
Qy	125	TTTCAAAATTCAGACAGATCTATGAAATTTGATATCATCTGTATGAGCCAGATGTAC	184
Db	368	TTTCAAAATTCAGACAGATCTATGAAATTTGATATCATCTGTATGAGCCAGATGTAC	427
Qy	185	TGCTATGACTTCACTTCTGAGACATTTACTGCTCATGATTTCCAGATTTGCA	244
Db	428	TGCTATGACTTCACTTCTGAGACATTTACTGCTCATGATTTCCAGATTTGCA	487
Qy	245	AATGGCAGAGTGCACCCCTCTGCTCTTGAAGCGAAATGAAAGTACGCCCCAG	304
Db	488	AATGGCAGAGTGCACCCCTCTGCTCTTGAAGCGAAATGAAAGTACGCCCCAG	547
Qy	305	AGAAATTTGTAGACATCAAGAAACTTTGGAAGAGAGACTGCGCAGGCTCTG	364
Db	548	AGAAATTTGTAGACATCAAGAAACTTTGGAAGAGAGACTGCGCAGGCTCTG	607
Qy	365	TGATCTGGACTGACTGTGATAGAGAGCGAAACATCGGCTTTGAGATTATCCAGCTGT	424
Db	608	TGATCTGGACTGACTGTGATAGAGAGCGAAACATCGGCTTTGAGATTATCCAGCTGT	667
Qy	425	GTAAGGCTGTAAAGCCCAATGTGACGNGTGCACGCCGGAATTCGTGAGATCAACCCC	484
Db	668	GTAAGGCTGTAAAGCCCAATGTGACGNGTGCACGCCGGAATTCGTGAGATCAACCCC	727
Qy	485	ATGCGCTCAGGACAGACTTGTGAAAAACCTTGACGAGCCGTGATCAGAGGGTGAAGCTGTG	544
Db	728	ATGCGCTCAGGACAGACTTGTGAAAAACCTTGACGAGCCGTGATCAGAGGGTGAAGCTGTG	787
Qy	545	TGATGTAGGACGAGAGCTGAGACTGAGAGATTGGAGCTGCTTACTAGTTCAGAGACC	604
Db	788	TGATGTAGGACGAGAGCTGAGACTGAGAGATTGGAGCTGCTTACTAGTTCAGAGACC	847
Qy	605	TGCGGCTTCAGAGGATTTTCTGAGGTGCTGGCAGAGACACTCATCGTTACGCGAGCT	664
Db	848	TGCGGCTTCAGAGGATTTTCTGAGGTGCTGGCAGAGACACTCATCGTTACGCGAGCT	907
Qy	665	GCCAGTTCCCCACACATGCGCTTGTGTGAGCGGTTCAAGCCATTGAGGCTTTGTAC	724
Db	908	GCCAGTTCCCCACACATGCGCTTGTGTGAGCGGTTCAAGCCATTGAGGCTTTGTAC	967
Qy	725	CAGAAATCTTCACAGAAATTAAGTAAGTCACTGACACCAAAAGATGATCGTACAAATTC	784
Db	968	CAGAAATCTTCACAGAAATTAAGTAAGTCACTGACACCAAAAGATGATCGTACAAATTC	1027
Qy	785	ACTGAAAAAGGACATCGACTTTTAACACACAGCGGTGCTGATCATGAGTTGTG	844
Db	1028	ACTGAAAAAGGACATCGACTTTTAACACACAGCGGTGCTGATCATGAGTTGTG	1087
Qy	845	TGGAGGATCCCATGGCACTGTGTAGAGGTCAAGATCAAGCCCAAGAGCAAGTGGCGGC	904
Db	1088	TGGAGGATCCCATGGCACTGTGTAGAGGTCAAGATCAAGCCCAAGAGCAAGTGGCGGC	1147
Qy	905	CTCAAGCCTTGAGACACTGTGAGGCTTGAGAACTGCGCTTCGAAAGTTGAAATTAATG	964
Db	1148	CTCAAGCCTTGAGACACTGTGAGGCTTGAGAACTGCGCTTCGAAAGTTGAAATTAATG	1207
Qy	965	CTAAGAAGACCATGAGATTGCTGAGAAAGCTTCACGTCAAGGGGTACATCGCTATATCCC	1024
Db	1208	CTAAGAAGACCATGAGATTGCTGAGAAAGCTTCACGTCAAGGGGTACATCGCTATATCCC	1267
Qy	1025	GAACAGAAACAAACATTTTCCAGAGACTTAAACCTGACGGTGTGTGGAACAGCAGA	1084
Db	1268	GAACAGAAACAAACATTTTCCAGAGACTTAAACCTGACGGTGTGTGGAACAGCAGA	1327
Qy	1085	CCCCGATCCACGCTGGGGGGCTTTGGCCAGAGCAATTTAGAGCGGGGTGTCCACACC	1144

```

|||||
Db 1328 CCCCGATCAGCGTGGGGGGCTTTGGCCAGAGCATCTAGAGCGGGGGTGGCCACCC 1387
Oy 1145 CACGCAATGGGAACAAGTGTGACCAAGCTCACCCTCCCATATCCACCACCAATACACA 1204
Db 1388 CACGCAATGGGAACAAGTGTGACCAAGCTCACCCTCCCATATCCACCACCAATACACA 1447
Oy 1205 ACAACTTACAGAGAGATGAAACAGCCACTGTACAGATTTATTTGTCGCATTTCTGGCTT 1264
Db 1448 ACAACTTACAGAGAGATGAAACAGCCACTGTACAGATTTATTTGTCGCATTTCTGGCTT 1507
Oy 1265 GCTGCTCCAGAGATGCTCAAGGGGAGAGAGACACACAGTGGAGATGCACATCGCTCAGAAC 1324
Db 1508 GCTGCTCCAGAGATGCTCAAGGGGAGAGAGACACACAGTGGAGATGCACATCGCTCAGAAC 1567
Oy 1325 GCTTTTGGCCCATGGCTCTCATGATTTCTGGCCGAAACTATCTGAGTGTGTATCCATPAG 1384
Db 1568 GCTTTTGGCCCATGGCTCTCATGATTTCTGGCCGAAACTATCTGAGTGTGTATCCATATG 1627
Oy 1385 ATCACTGAGAGTGAAGATCTCTCTGTATGAGCAAGAGATCCACTTTCAGCCACAGA 1444
Db 1628 ATCACTGAGAGTGAAGATCTCTCTGTATGAGCAAGAGATCCACTTTCAGCCACAGA 1687
Oy 1445 CCGTGGAGATGTTGAGACGGGGAGACCCACCCAGCTGCTCACCGAGGCGGACCTCA 1504
Db 1688 CCGTGGAGATGTTGAGACGGGGAGACCCACCCAGCTGCTCACCGAGGCGGACCTCA 1747
Oy 1505 TTGGCCCTCATGAGAAAGCATGGCATTTGTACGGATGCCACTCATCGCGAGACATGAGA 1564
Db 1748 TTGGCCCTCATGAGAAAGCATGGCATTTGTACGGATGCCACTCATCGCGAGACATGAGA 1807
Oy 1565 CCATCAAAAGCCGGAGTGTAGTGGGCTCACCCAGACAAAGGGTTCCTCCCTGGGCGAC 1624
Db 1808 CCATCAAAAGCCGGAGTGTAGTGGGCTCACCCAGACAAAGGGTTCCTCCCTGGGCGAC 1867
Oy 1625 TGGGCTGAGACTTGTGGAAGGTTATGATTCATGGGCTATGAATGTCTAAACCTGACC 1684
Db 1868 TGGGCTGAGACTTGTGGAAGGTTATGATTCATGGGCTATGAATGTCTAAACCTGACC 1927
Oy 1685 TCGGGGCTCAACTGGAAGCTGATCTGATGATCTGTGATGGCAAAAGGACAAATTTG 1744
Db 1928 TCGGGGCTCAACTGGAAGCTGATCTGATGATCTGTGATGGCAAAAGGACAAATTTG 1987
Oy 1745 TGGTTTAAAGCGAGCAAGTGCAGAAATACAGAGGTTTTCATTGAAAGCGGTGGCTAAAG 1804
Db 1988 TGGTTTAAAGCGAGCAAGTGCAGAAATACAGAGGTTTTCATTGAAAGCGGTGGCTAAAG 2047
Oy 1805 CAAAGAAATTGACGAGGCTTGGCCAGTACTTTGGGAATGGGAGACAGATTGGCCAGC 1864
Db 2048 CAAAGAAATTGACGAGGCTTGGCCAGTACTTTGGGAATGGGAGACAGATTGGCCAGC 2107
Oy 1865 AAGAGATATCTACCAAGCCATGCCAGAGCCCATCAGAAAGTCCCAACAGTGCACAGG 1924
Db 2108 AAGAGATATCTACCAAGCCATGCCAGAGCCCATCAGAAAGTCCCAACAGTGCACAGG 2167
Oy 1925 ACATGTCCTTAAAGCAAGAAATGGGGGTTTACCTACCTGATGAGGTTTCCAG 1984
Db 2168 ACATGTCCTTAAAGCAAGAAATGGGGGTTTACCTACCTGATGAGGTTTCCAG 2227
Oy 1985 AGTGTGCTCAGCTGTGTGGCTTCTGATCTCGGTGCTGAGGCGCAGCAGGAGACAGATG 2044
Db 2228 AGTGTGCTCAGCTGTGTGGCTTCTGATCTCGGTGCTGAGGCGCAGCAGGAGACAGATG 2287
Oy 2045 TGTGTCACTTTGTGACCAACACCCCTGTGTACAGGTTTAAAGTTTAAAGTTCAGCGGTA 2104
Db 2288 TGTGTCACTTTGTGACCAACACCCCTGTGTACAGGTTTAAAGTTTAAAGTTCAGCGGTA 2347
Oy 2105 GCCTTCCCGGACATGCTGAGATTTGTTGTCGATCGCGGATGAGACAGCC 2164
Db 2348 GCCTTCCCGGACATGCTGAGATTTGTTGTCGATCGCGGATGAGACAGCC 2407
Oy 2165 TGAGGAGATCTTGACCTGAGATTTTTCAGGGGGCCCCCAGAGGCTAGCCAGCCCTCTG 2224
|||||

```

```

Db 2408 TGAGGAGATCTTGACACTGATGATTTTTCAGGGGGGGCCCCCAGAGGCTAGCAGACCCCTTG 2467
Oy 2225 GCCGCTGCAAGGTAAACCAAGTCCCTTAACAGAGTGGACAAACAGCCAGCACCCTGACCTG 2284
Db 2468 GCCGCTGCAAGGTAAACCAAGTCCCTTAACAGAGTGGACAAACAGCCAGCACCCTGACCTG 2527
Oy 2285 CTGACAGCAGACGACTGGGTCTCAAGAGCTCTGGCCCAAGACCCCTCCACACCCAGCG 2344
Db 2528 CTGACAGCAGACGACTGGGTCTCAAGAGCTCTGGCCCAAGACCCCTCCACACCCAGCG 2587
Oy 2345 CTGCTGTGAAMACATTTCTGTACCTGCACATGTGGCCAGAGAGCTGTCTGCTCACTG 2404
Db 2588 CTGCTGTGAAMACATTTCTGTACCTGCACATGTGGCCAGAGAGCTGTCTGCTCACTG 2647
Oy 2405 TCCGTAAGGAGGGCCCCAACCGGGGCGGAGTTCCTTTAAAGTCAACGAGAGTACTGCA 2464
Db 2648 TCCGTAAGGAGGGCCCCAACCGGGGCGGAGTTCCTTTAAAGTCAACGAGAGTACTGCA 2707
Oy 2465 ACTTCTTCCTGTGGGAGAGAGAGCCCAATCCGGGAGCAGAGAGGGGCTCCGCTTGGCAT 2524
Db 2708 ACTTCTTCCTGTGGGAGAGAGAGCCCAATCCGGGAGCAGAGAGGGGCTCCGCTTGGCAT 2767
Oy 2525 ATGACCCCTGGGCGCTCCCTGATGGATGCCACACAGGCGCCAGGATCCACTAGTGGGT 2584
Db 2768 ATGACCCCTGGGCGCTCCCTGATGGATGCCACACAGGCGCCAGGATCCACTAGTGGGT 2827
Oy 2585 TTGGCAACCTGCTGTATGGCAGTGTGATGGACATCTGCTCTTTCAGAGCCACCTCCG 2644
Db 2828 TTGGCAACCTGCTGTATGGCAGTGTGATGGACATCTGCTCTTTCAGAGCCACCTCCG 2887
Oy 2645 TCACACGAGCTGTGAGAAAGATGAGACCAACAAAGGGGGGCGCGTTCACACATGTGCCA 2704
Db 2888 TCACACGAGCTGTGAGAAAGATGAGACCAACAAAGGGGGGCGCGTTCACACATGTGCCA 2947
Oy 2705 AGCCGAGAGAGCAGCAGTGTGCTTTTCCAGTGGGTGCATGAGAAACACCGCTCCAGGGA 2764
Db 2948 AGCCGAGAGAGCAGCAGTGTGCTTTTCCAGTGGGTGCATGAGAAACACCGCTCCAGGGA 3007
Oy 2765 CTTCCTGAGAGCCCGTCTCTGAGACAGAGACAGAGAAACCTGTGAGTGGGAAGCCAGAA 2824
Db 3008 CTTCCTGAGAGCCCGTCTCTGAGACAGAGACAGAGAAACCTGTGAGTGGGAAGCCAGAA 3067
Oy 2825 GCAAAAGCCCGCGGCGCAGTCTCTCAGACATGGGGTCCACAGCAAGAAACCCCGGAAT 2884
Db 3068 GCAAAAGCCCGCGGCGCAGTCTCTCAGACATGGGGTCCACAGCAAGAAACCCCGGAAT 3127
Oy 2885 GCAGCCTTTGCCACACAGCCTGAGACACACCCGTCCTTTTGTCTCTCAGAACAGATGAGCTC 2944
Db 3128 GCAGCCTTTGCCACACAGCCTGAGACACACCCGTCCTTTTGTCTCTCAGAACAGATGAGCTC 3187
Oy 2945 AGGCTAGGTTAGGAA 2960
Db 3188 AGGCTAGGTTAGGAA 3203

RESULT 2
AB006074 3741 bp mRNA linear ROD 18-APR-1998
LOCUS AB006074
DEFINITION Mus musculus mTOR3 mRNA for topoisomerase III, complete cds.
ACCESSION AB006074
VERSION AB006074.1 GI:3061307
KEYWORDS mTOR3; topoisomerase III.
SOURCE Mus musculus (strain: Balb/c) testis cDNA to mRNA.
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 3741)
AUTHORS Seki,T.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-1997) Takahiko Seki, Tohoku University, Faculty
of Pharmaceutical Sciences, Aoba Aramaki, Aoba-ku, Sendai, Miyagi
980-77, Japan (E-mail: takahima12.pharm.tohoku.ac.jp,
Tel: +81-22-217-6876, Fax: +81-22-217-6873)

```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
2 (sites)	Seki,T., Seki,M., Katada,T. and Enomoto,T.	Isolation of a cDNA encoding mouse DNA topoisomerase III which is highly expressed at the mRNA level in the testis	Biochim. Biophys. Acta 1396 (2), 127-131 (1998)	98201702	
		Sequence	updated (02-Oct-1997).		
		Location/Qualifiers			
		1. 3741			
		/organism="Mus musculus"			
		/strain="Balb/c"			
		/db_xref="taxon:10090"			
		/tissue_type="testis"			
		266..3277			
		/gene="mtop3"			
		266..3277			
		/gene="mtop3"			
		/codon_start=1			
		/product="topoisomerase III"			
		/protein_id="BAA25662.1"			
		/db_xref="gi:3061308"			
		/translation="MIPPVTLFAFMHRRPGRALSRAMEVAREGVKLYLCAEKNIDKAAKIDILNSGRMRKRGESLKNKITYEPDHLGQNWMTLVSSHLAHDEQMOFRKMSQNDLYLFEALIKGCYDNPENIDIKLTPEDHNSQALVITMSDHLADRGENTFEIITHCKAVPNLRVLRARSETTPPAVTPACDNLTPEORSDADVAKOEDLRIGAFTRQPTLRQRIEPEVLAEQLISTSGSCPTLGEVVEFKAIQAEPVEVFKIKYTHDKKDQVEFNMKRYRLFNHTACLVLQLOCMEDPMANVVRSPKSKMRPOALDIYELRKLASKRLINAKETRIAEKLYTQGISYSPRETIKPKDLNLVALVROGVDPHMGAFNQTLLERGGPTPRNGSKSDQAHNP1HPTKTSGLGDQDRLYEIVHPLACSGDOAGQETVEIDIAOEFVAGHGLILARANLDVYPYHMSDRLILPYEGSHQSPRVEAVDGETSPPOLIDAEADILALMERHGGTTLTAHNEIETKARYVELITSDKRLPHLDMGLVGVDSMGTEMSKPDRLRELEADLILCEGKDKOVLRQOVKTQVTEAVNARAKIDEALSQYLGERTMAOQEEIYPAMEPEVRKCPQCNKDMVLTKRSGGVYLSMCEFPKCSAAWVPSQVLYEASRDNSVSCQPPVYRLKLFKRGSLPEPAMLEFVGCIGCCDETLEKLEIFGLRFPALPRASQSGHLLQASOLNMDSSOHNLSOPLVNRHTREKTYVAQALLPPTAGESNSVTNCGREAVLLIVRKQRPQGRHFFKCSMGDNFLPMDSSHS TGSGTTPSASGPRESSVSGCRSSVSIMHDGRLSDSDGGTPELCGPATYTRYOKGSPNRQFHTCAKPREQCGFQWVDVAVESFAAMPVGGRGAKQRPDEASKRPBAGSSDSSTVKKRRKCSLIHOHPHTRFPCPNR"			
		3741			
		/note="35 a nucleotides"			
		BASE COUNT 925 a 954 c 1048 g 814 t			
		ORIGIN			
		Query Match 75.1%; Score 2233.2; DB 10; Length 3741;			
		Best Local Similarity 85.0%; Pred. No. 0;			
		Matches 2512; Conservative 0; Mismatches 438; Indels 6; Gaps 1;			
OY	11	TCGAGATGGCCCTCGGAGCGTGCGGAAAGTCTCTGTGTGGCCGAAAAACGACGCGG	70		
	1				
Db	342	TGGAAGTGGCCTTCGAGAGAGGCGAAAAAGTTCTCTCGGTGGCGGAGAAAAACGACGCTG	401		
OY	71	CCAAGGGATGGCGACCTGCTGTCAAACGGTCCGATGAGCGGAGAGAAAGACTTTCAA	130		
	1				
Db	402	CCAAAGGGATGGCGACCTGCTGTCAAACGGGCGTATAGGGGAAAAAGAGCCCTTTCTTA	461		
OY	131	AATTTAAAGAAGATCATTAATTTGATATCATCTGTATGCGCAGAAATTTACATGTAA	190		
	1				
Db	462	AATTTAAAGAAGATTTATGAATTTGACTATCATCTGTATAGCCAGAAATTTACTATGATAA	521		
OY	191	TGACTTCAGTTTCTGGACATTTACTGSGCTCATGTATTCGACATGCGATTTGAAAAATGGC	250		
	1				
Db	522	TGACTTCAGTCTCTGGAATTTGCTGCTCATGATCTCAAGATGAGATTTGGAAGATGGC	581		
OY	251	AGAGCTGGAACCCCTGTCTCTCTTTGAAAGCAAAATTTGAAAAAGTACGCCAGAGATT	310		
	1				
Db	562	AGAGCTGGAATCCCTTCTCTCTTTGGAAGCAAAATTTGAAAAAGTATTTGCCAGAAAAATT	641		
OY	311	TTGTGAGATCATGAAGAACTTTGGAAGACAGAGACTCGCGACGAGCTGTGGTATCT	370		
	1				
Db	642	TTATGAGACATCAAGAAACTCTAGACAGAGACACATCATTTGTAGGCCCTGGTGATCT	701		
OY	371	GGACTGACTGTGATGAGAGGCGAAAAATCATCGGGTTTGAGATTATTCACAGCTGTGTAAAG	430		

[illegible]

Db	1782	AAATGTGTGATGGAGAGACTAGACCCACCCCACTGCTTACCGAAGCTGACCTCATCGCC	1841
Qy	1511	TCATGGAGAAAGCATGGCATTTGGTACGGATGGCCACTCATGCGGAGCACAATCGAGACATCA	1570
Db	1842	TCATGTAGAAACATGGGATGGGTAGTGTAGTGAACATCATGACAAACATTTAGAGACATCA	1901
Qy	1571	AAGCCGGATGTACGTGGGCTCAACCCAGACAGCGGTTCTCCCTGGGCACTGGGCA	1630
Db	1902	AAGCCGGATGTATGTGGGACTCACTCACTCAGACAGAGGGGTCTCGCGGGGCACTGGGCA	1961
Qy	1631	TGGGACTTGTGGAAAGTTATGATTCATGGGGCATGAAATGTCTAAAGCTCAACCTCCGG	1690
Db	1962	TGGGACTTGTAGAAAGTTATGATTCATGGGGTATGATGATGATCAAGGCTGACCTCCGTG	2021
Qy	1691	CTGAACCTGGAAGCGATCTGAAGCTCATGCTGTATGGCAAAAAGACAAATTTGTGGTTTC	1750
Db	2022	CCGAGCTGGAAGCGTGAATCTGAAGCTATCTGTATGAGGCAAGAAAGATTAAGTTTCAAGTTC	2081
Qy	1751	TAAAGCAGCAAGTGCAGAAATACAGCAGGTTTTCATTGAAGCGGTGGCTAAAGCAAGA	1810
Db	2082	TAAAGCAGCAAGTGCAGAAATACAGCAGGTTTTCATTGAACACTGGCGCAAGGCAAGA	2141
Qy	1811	AATGGAGCGAGGGGCTTGCGCCACTACTTTTGGGAATGGGACAAGTTGGCCCAAGAAAG	1870
Db	2142	AGTTGGATGAGGGCTTTATCTCAATCTACTTATGAGAAAGAAACAGATGGGCCCAAGAAAG	2201
Qy	1871	ATATCTACCCAGCATCCAGAGCCCATCAGAAAGTGGCCCAAGTGCAAACAGCATAGG	1930
Db	2202	AGATCTACCCAGCATCCAGAGCCCTGTTCGGAAAGTGGCCCAAGTGTAAACAGGATATGG	2261
Qy	1931	TCCTTAAAGACCAAGAAAGATGGGGGTTCTACCTCACTGATGGGTTTCCAGAGTTC	1990
Db	2262	TCCTCAAGACCAAGAAAGTGGGGTTCTACCTCACTGATGGGTTTCCAGAGTTC	2321
Qy	1991	GCTCAGCTGTGTGGCTTCCTGACTCGGTGCTGGAGCCACAGCAGCACTGTGTCTC	2050
Db	2322	GGTGAGCTGTGTGGTTTCTGACTCGGTGCTGGAGCCACAGCAGCACTGTGTCTT	2381
Qy	2051	CAGTTTGTCAAGCCACACCTCGTGTACAGAGTTAAAGTTTAAAGCGGCTTACCTTC	2110
Db	2382	CTGTTTGTCAAGCCACACCTCGTGTACAGAGTTAAAGTTTAAAGCGTGTACCTTC	2441
Qy	2111	CCCCAGCATGCTCTGGAGTTTGTGTCTGCATCGCGCGATGCGACGACACCTCGAGG	2170
Db	2442	CCCCAGCATGCCCCCTGGAGTTTGTGGCGTGCATAGGTGATGTGATGAGACTTTGAAG	2501
Qy	2171	AGATCTGTGACCTTAGATTTTTCAGGGGGCCCCCAGGGCTTAGCAGCCCTGTGGCCCC	2230
Db	2502	AAATCTTGTGGGCTCGATTTTTCAGGGGCTCTCCCAAGACTTAGCAGCCCTGTGGCCCC	2561
Qy	2231	TGCAGGGTAACCACTGCTCCCTGAACAGGATGGACAAAGCCAGCACCCCGAG-----CTG	2284
Db	2562	TGCAGGGTAACCACTGCTCCCTGAACAGGATGGACAGCACTAGCAATTAACCTGTCAACTTC	2621
Qy	2285	CTGACAGCAGACAGACTGGGGTCTCTCAAAAGGCTTGGCCAGACCTTCCACACCCACAGG	2344
Db	2622	TGGTTAAACAGACATCAAGACCTTTCAAAAGACAGATGGCCCAAGCGTTGTACTACCCACTA	2681
Qy	2345	CTGCTGGTGAAGCAATTTCTGTGACCTGCACACTGTGGCCAGAGAGCTGTGTCTACTG	2404
Db	2682	CTGCTGGTGAAGCAATTTCTGTGACCTGTCAATTTGTGGCCGAGAGGCTGTGTCTCACTG	2741
Qy	2405	TCCGTAAGGAGGGCCCAACCGGGGGCGGAGTTCTTAAATGTGCAAGGAGAGGTACTACTCA	2464
Db	2742	TCCGTAAGGAGGGCCCAACCGGGGGCGGAGTTCTTCAAGTGCAGTAATGTGTACTCA	2801
Qy	2465	ACTTCTTCTGTGGGAGAGAGCCCAATCCGGAGACAGAGAGGCTCTCTCTGGCAT	2524
Db	2802	ACTTCTTCTGTGGGAGAGAGCCCAATCCGGAGACAGAGAGGCTCTCTCTGGCAT	2861
Qy	2525	ATAGACCCCTGGGCGCTCCCTGGGATGCCACAGGCCAGAGATCCACTTAAAGTGGGT	2584
Db	2862	CAGAGCCCCAGGCGCTGTGTAGAGAGCCCATTCAGTGTAGGACAGCCCATGATGATGGGT	2921

[illegible]

Query Match 23.1%; Score 686; DB 9; Length 132486;
Best Local Similarity 100.0%; Pred. No. 6,6e-172;
Matches 686; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2077	AGGTTAAAGTTAAAGTTAAAGCGGGGATGCTTCCCGGACCAATGCTCTGGAGTTTGT	2136
DB	53784	AGGTTAAAGTTAAAGTTAAAGCGGGGATGCTTCCCGGACCAATGCTCTGGAGTTTGT	53725
QY	2137	TGCGATCGCGCGGATGCGAGACACCCCTGAGGAGATCTCGACCTGAGATTTTCAGGG	2196
DB	53724	TGCGATCGCGCGGATGCGAGACACCCCTGAGGAGATCTCGACCTGAGATTTTCAGGG	53665
QY	2197	GGCCCCCGGAGGCTGACGAGCCCTGCGCCCTGACAGGCTAACAGTCCCTGACAG	2236
DB	53664	GGCCCCCGGAGGCTGACGAGCCCTGCGCCCTGACAGGCTAACAGTCCCTGACAG	53605
QY	2257	ATGACAAAGACCCGACACCCCTGCTGACAGCAGACAGTGGGTCTCTAAAGCT	2316
DB	53604	ATGACAAAGACCCGACACCCCTGCTGACAGCAGACAGTGGGTCTCTAAAGCT	53545
QY	2317	CTGCGCCAGACCTCCACACACCCGCTGCTGTAAGCAATTTCTGACCTGCAAC	2376
DB	53544	CTGCGCCAGACCTCCACACACCCGCTGCTGTAAGCAATTTCTGACCTGCAAC	53485
QY	2377	TGTGGCCAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2436
DB	53484	TGTGGCCAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	53425
QY	2437	TTCCTTAAAGTCAAGGAGGAGTGAACCTTCTCTGCTGCTGCTGCTGCTGCTGCTG	2496
DB	53424	TTCCTTAAAGTCAAGGAGGAGTGAACCTTCTCTGCTGCTGCTGCTGCTGCTGCTG	53365
QY	2497	GGAACAGAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2556
DB	53364	GGAACAGAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	53305
QY	2557	CGAGGCCAGAGGATCCACCTGAGTGGGTTTGGCAACCTGCTGCTGCTGCTGCTG	2616
DB	53304	CGAGGCCAGAGGATCCACCTGAGTGGGTTTGGCAACCTGCTGCTGCTGCTGCTG	53245
QY	2617	ACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2676
DB	53244	ACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	53185
QY	2677	AAGGGCGGCGAGTTCACACATGTCACCAAGCGAGAGAGCAGCAGTGTGCTTTTCAG	2736
DB	53184	AAGGGCGGCGAGTTCACACATGTCACCAAGCGAGAGAGCAGCAGTGTGCTTTTCAG	53125
QY	2737	TGGGTGATGAGAACACCGCTCCAGG	2762
DB	53124	TGGGTGATGAGAACACCGCTCCAGG	53099
RESULT 4			
LOCUS	AF255733	3891 bp	mRNA linear INV 21-MAY-2000
DEFINITION	Drosophila melanogaster DNA topoisomerase III alpha (CG10123) mRNA,		
ACCESSION	AF255733		
VERSION	AF255733.1	GI:7960301	
KEYWORDS			
SOURCE	fruit fly.		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
	pterygota; Neoptera; Endopterygota; Diptera; Brachyera;		
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 3891)		
AUTHORS	Plank,J.L., Reineke,J.C., Wilson,T.M. and Hsieh,T.-S.		
TITLE	Drosophila melanogaster topoisomerase III alpha		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 3891)		
AUTHORS	Plank,J.L., Reineke,J.C., Wilson,T.M. and Hsieh,T.-S.		

TITLE Direct Submission
JOURNAL Submitted (13-APR-2000) Biochemistry, Duke University Medical Center, Research Drive, Durham, NC 27710, USA
FEATURES
Source location/Qualifiers
1.3891

gene
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2"
/map="2L: 37E3"
/dev_stage="embryo"
1.3891
/gene="CG10123"
/note="top3a"
37.3789
/gene="CG10123"
/codon_start=1
/product="DNA topoisomerase III alpha"
/protein_id="AAE71288.1"
/db_xref="GI:7960302"

CDS

/translation="MKAVLTCPRPVVAASHRRYCANPQANKYLINAEKNDKATNAGL
LSNGAORREGYSVYKVPEDFEAPVRGONAKMYTSSVGHMOLAOVSXKMYTDP
RSLFDPVKEVGSDEYIKRLTEREVGCGGLITWDCRGENIGETIIVDCRAIX
PNISVYRATSEITTVAVRRALQOLGPOKROSDAVDRFELDLPTGAATIRQTRRL
QRFPERKIDKLITVSCQIPRTGPAERKEIETEARVSEPFKIKYLIHTIDLTVEFN
WARKLDRKACENYLLCLAEPPRALVESYVKKRKHRTPLDTVEKLSKRL
KLSAKETMTAEKLYTKGFTSYRTEINQSKFALAPLVEMOTGHRDGAERQVIE
WGPNPNGKNSDQADHPHPTKLAENLOGNEARVIELVYHFTLACYSKAVGSETLVH
IDLAGEKFTANGLVIERNYLDVYVYDGNANAKOIHHEMORPREPEVLSIGETAP
PLTLEADLILMEKHGIGDPAHAENHNTKEKGYIGLDKRGVIGMGLYEGED
AMELALAKPOLRAEPFLDKLIGOGKROGPDVFTDILAKYKAOYDITPKITAMDKI
SARINTPANAANSVQGDGDSASRSHITQSIPOCPKCNAPRALRKKNQOCYATCN
NFPDCKNAWMLPTECKDASVLDCECTCGDGTMLFRJSTPIYKGFPGPSWNTIC
LPQDNLFRFTFININDSVKVGIVGIVEGKGGLGPGPBGSGSGAGSGSGSPG
GGGSGGGSG
KPKRTKEPRAAPKRTSSKSGSIRSEFTSAATNSAGNLDEFSDNGFEDAMLEA
AESVSSSOPKRTISVPLDDDDIAAFAADDAEFALVNGTPTESNDQDOLKSL
EYVLCCTGCOQPARONTVRKNGPMLGLYIKCPKPDCCNFQANDPSSAKSKNSG
APQSTTSMNSRVYTLPSIQOINSORGOSSMRSNSSYVITQTKRQDERNATNGD
GSEVMNCQQLASQLTVRMDGPNMGPFYACPTREKSGCFKMGKDQNOGASSTSG
SANRNPQCSQPAITSDGPKTRRCGLCKRKHRTKPKCRKDEFDM"

BASE COUNT 1030 a 973 c 1095 g 793 t
ORIGIN

Query Match 20.0%; Score 594.6; DB 3; Length 3891;
Best Local Similarity 59.6%; Pred. No. 1.1e-147;
Matches 1043; Conservative 0; Mismatches 699; Indels 9; Gaps 2;

QY	36	GAAAGTCTCTGTGTGGCCCAAAAAACGACGCGCCCAAGGGATGCGGACCTGCTGTC	95
DB	114	GAAAGTCTCTGTGTGGCCCAAAAAACGACGCGCCCAAGGGATGCGGACCTGCTGTC	173
QY	96	AAACGATCGCATGAGCGGAGAGAGACCTTCAAAATTCACAAGATCTATGAATTTGA	155
DB	174	AAATGCGCGACGACCAAGGCGCGGAGGCTACTCGGTATACACAAGGTCTTCGATTGCA	233
QY	156	TTATCATCTGTATGCGCAAGATGTTACATGTAATGATGATCTTCACTTCTGACATTTACT	215
DB	234	GGCCCTGTGCGGCGTCAAGATGTCACAAATGTTATGACCTCCGTATCCGGCCACATGAT	293
QY	216	GGCTCATGATTTCCAGATGACAGTTTCGAAAAATGGCAAGCTGCACACCTCTTGCTCTT	275
DB	294	GCAAGTGGCATTTCCAGGTGTCTTCAAAAATTTGAGAGAGGTGGATCCGGCTTTATT	353
QY	276	TGAAGCAGAAATTTGAAGTACTGCCAGAAATTTTGTAGACATCAAGAAACTTTTGA	335
DB	354	CGACGCGCGCGGTGAAAAGGCGGTGGGCTGGGATTTATGACCTTATCAAAAGACCTTTGGA	413
QY	336	ACGAGAGACTCGCCAGTGCAGGCTGTGTGATCTGACCTGACTGTGATAGAGAAGCGCA	395
DB	414	ACGGGAGGTGCGGCGGCTGTCAGGATTTATCTGACGAGATTTGGCATCGCGAGGCGCA	473
QY	396	AAACATCGGGCTTTGAGATTTATCCACGCTGTAAAGCTGTAAAGCCCAATCTGCAGGTGTT	455

Db	474	GAACATTGGCTACGAGATTATGATGTGTGCGCTATCAACCAATATTTGCGTTTA	533
Qy	456	GGAGGCCCATTTCTCTGAGATCAGACCCATCCGCTCAGAGACGCTTGTGAAGAACTGAC	515
Db	534	TGCTGACACTTCTTCGAGATTACCAAGGTGGTGGCTCGGCTCTGTGAGCAATTTGG	593
Qy	516	CGAGCCTGATCAGAGGTGAGTGTGATGTGAGGACGACCTGTGAGACTGTGAGAT	575
Db	594	GCAGCCGAGCAAGAGCAGGCGATGAGTGTGCGACGGAAGTGTGATTTGCGAAC	653
Qy	576	TGAGAGTGGCTTACTAGTTCCAGACCCGCGCTTCAGAGATTTTCTTGAGGTGCT	635
Db	654	TGCTGCGCCATACCCCGGTTTCAGCAANTGCCCTACAAAGCTGTCTTCCCGCAAGAA	713
Qy	636	GCGAGACAGCTCATGATTTAGGCGACCTGCGCATTTCCACACTGCGCTTGTGTGGA	695
Db	714	AGCTGATTAAGCTCATATTCATGGAAGCTGCGCAAAATTCACCTGCGATTCGCGCGA	773
Qy	696	GCGCTCAAGACCATTCAGGCTTTTGTACAGAAATCTTCCACAGAAATTAAGTAACTCA	755
Db	774	ACGCTACAAAGAGATGAGCGCTTGTATCCGAGCCATTTTGGAAAGATTAAGGTCTGCA	833
Qy	756	TGACACAAAGATGTATGCTAGATTTCAACTGAAAGGATCGACTCTTTAACCAAC	815
Db	834	CACAAATGATGACTTACGATGCGTGTGATTTAATTTGGCCCGCAACCGACTCTTCATTAAGGA	893
Qy	816	GCGTTGCTAGTTCTCTATCAGTTGTGTGTG-----AGATCCCATGGCAACTGTGTG	869
Db	894	AGCTTGTGAAGAACTATTTGCTCTGCTGCGCGGCGAACCAGATCTAGAGCGGTGGA	953
Qy	870	AGAGCTCAGATCTAAGCCCAAGAGCAAGTGGCGCTCAAGCTTGTGACACTGTGAGCT	929
Db	954	GAGCGTACCGCTTAACCCCAACAAATGCGGCCCAACTCCGCTGATTAACCTGTGAAT	1013
Qy	930	TGAGAGCTGGCTTCTCGAAGTTGAAATTAATGCTAAAGAAACATGAGATTTGCTGA	989
Db	1014	GGAGAAAGCTTGGCTCGAGAAACTCAAGCATATGTGCCAAGAGAACATGACATAGCCGA	1073
Qy	990	GAGCTCTACACTCAAGGGTACATCAGTATCCCGCAAGCAAGAAACAATTTTCCAG	1049
Db	1074	AAATATATATACCAAGGGTTTCACTAGTATCCCGTACGAAACCAACAGATTTCTCA	1133
Qy	1050	AGACTTAACCTGAGCGGTGTGTGTGGAACAGACACCCCGATCCACGCTGGGGGGCTT	1109
Db	1134	GGAGTTCGCTGTGGACCACTAGTGTAGATGAAGGGGTCAACCGGACTCGGAGTGGACAT	1193
Qy	1110	TGCGCAGAGCATTTAGAGGGGGGTGTCACCCACGCAAGCAATGGAACAAAGTCTGACA	1169
Db	1194	TGCGCAGCGGGTATGAGTGG-----GGACCAAAATCTCTCGAAAGGTAAACAAATCGGATCA	1250
Qy	1170	AGCTCACTCCCATTCACCCACCAAAATACACCAACCAACTTACAGGAGATGAGACGC	1229
Db	1251	GGCCCATCTCCCAATTCATCCACCAAGCTGGCAGAAACCTCAGGAGAAACGAGCGCG	1310
Qy	1230	ACTGTACAGATTATTTGTGTGCGCATTTCTGCGTGTCTGTCTCCAGATGCTCAGGGCA	1289
Db	1311	AGGTACAGAACTGTGTCTCGGACATTTCTGCGTGTCTGTGAGAAAGATGCACTGCGCTC	1370
Qy	1290	GGAGACACAGTGGAGATGACATCGCTCAGAGACCTTTGTGGCCATGGCTCATGAT	1349
Db	1371	CGAAACCTTTGTCCATATGACATTTGCGGGTGAAGATTCACGGCCCAATGTGTGTTAT	1430
Qy	1350	TCTGGCCGCAAACTATCTGATGTGTATCCATATGATCTAGTGAAGTACAGATCTCTCC	1409
Db	1431	TCACGACGAAACTACTATGATTTATTTATTTATGACAACTGAGTGCACCAAGACATCA	1490
Qy	1410	TGCTATGAGCAAGATCCCACTTTCAGCCGACACCGTGTGAGATGTGTGACGGGAGAC	1469
Db	1491	TCACTATGAAAGCGGCGGTTTGAAGCCAGGAAGTGTCTCCACCAAGTGTCCAC	1550
Qy	1470	CACCCACCAACCTGCTCACGAGCGCAGCACTATTTGCCCTCATGTGAGAAAGATGGCAT	1529

Db	1551	CACTGCCCGCCCGCTGTGACCGAGCGGATTTGATTGACATGATGAAGAAAGCATGGAT	1610
Qy	1530	TGCTAGGATGCCATCTATGCGAGACATCGAGACCATCAAGCCCGATGACGTGG	1589
Db	1611	TGCTACAGATGCCACGACGCTGAGCATTAACACAGATCAAGAGCGCGGATCATCGG	1670
Qy	1590	CCTCACCCCGACAGACGGTCTCTCCCTGGGACGCTGGGATGGGACTGTGGAAGTTA	1649
Db	1671	AGTGTGATTAAGGCGCTTTTGTGTACCTGATGATTTATTTGATGAGTGTATGAGGATA	1730
Qy	1650	TGATTCATGAGGCTATGAATGCTAAGCTGACCTCCGCGCTGAGATGGAAGCTGATCT	1709
Db	1731	CAATGCGATGAGACTGTGCGCTTGCACAGCAACACTCTGCTGATGATTTGAATTTGAT	1790
Qy	1710	GAACTGATCTGTGATGCGCAAGAAAGACAAATTTGTGTTCTAAGCAGCAGCACTGCAAA	1769
Db	1791	GAACTAATTTTGCAGCGCCAGAAAGATCCCAAGGTGTCTGATCCGCAACAGATTAACAA	1850
Qy	1770	ATPACAAGCAGG	1780
Db	1851	GTACAAGCAGG	1861

RESULT 5	AC097545/c					
LOCUS	AC097545	88718 bp	DNA	linear	HTG 20-DEC-2001	
DEFINITION	Rattus norvegicus clone CH230-6BE7, *** SEQUENCING IN PROGRESS ***.					
ACCESSION	AC097545					
VERSION	AC097545.3					
KEYWORDS	HTG: HTGS, PHASEL.					
SOURCE	Norway rat.					
ORGANISM	Rattus norvegicus					
	Eumariota, Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
REFERENCE	1 (bases 1 to 88718)					
AUTHORS	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsdrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbata, J., Benton, J., Blinze, K., Blankenburg, K., Bonin, D., Bouch, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Bunay, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsli, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C., Kralovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Maubley, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, S., Nguyen, N., Nickerson, E., Nwokwenwo, S., Ogun, M., Okunolu, G., Oragunye, N., Oragunye, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Prims, E., Pu, L.L., Ruites, M., Ren, Y., Rives, M., Rojas, A., Rojupokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vanson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,					

TITLE
JOURNAL
AUTHORS
REFERENCE
JOURNAL
COMMENT

Watlington,S., Williams,G., Williamson,A., Wlaczek,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Unpublished
Direct Submission
2 (bases 1 to 88718)
Moriy,K.C.
Submitted (19-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17064400.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHY0
Center clone name: CH230-68E7
----- Summary Statistics
findthraplist
Assembly program: Phrap; version 0.990329first call to
Consensus quality: 74659 bases at least Q40
Consensus quality: 79368 bases at least Q30
Estimated insert size: 66722; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 0.8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 47 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
4658: contig of 4658 bp in length
4659: gap of unknown length
4759: contig of 3152 bp in length
7911: gap of unknown length
8010: gap of unknown length
11828: contig of 3818 bp in length
11829: gap of unknown length
15313: contig of 3385 bp in length
15314: gap of unknown length
15414: contig of 3126 bp in length
18539: gap of unknown length
18540: gap of unknown length
21828: contig of 2643 bp in length
21883: gap of unknown length
21883: gap of unknown length
24340: contig of 2958 bp in length
24341: gap of unknown length
24441: gap of unknown length
26396: contig of 1956 bp in length
26397: gap of unknown length
26496: gap of unknown length
29010: contig of 2514 bp in length
29011: gap of unknown length
30510: contig of 1400 bp in length
30511: gap of unknown length
30610: gap of unknown length
32919: contig of 2309 bp in length
33019: gap of unknown length
33020: gap of unknown length
34735: contig of 1716 bp in length
34736: gap of unknown length
34835: gap of unknown length
36525: contig of 1690 bp in length
36526: gap of unknown length
36625: gap of unknown length
38856: contig of 2231 bp in length
38857: gap of unknown length
38956: gap of unknown length
40235: contig of 1279 bp in length
40236: gap of unknown length
41823: contig of 1488 bp in length
41824: gap of unknown length
41923: gap of unknown length
43349: contig of 1426 bp in length

43350 43449: gap of unknown length
43450 45217: contig of 1768 bp in length
45218 45317: gap of unknown length
45318 47327: contig of 2010 bp in length
47328 47427: gap of unknown length
47428 49395: contig of 1968 bp in length
49396 49495: gap of unknown length
49496 51028: contig of 1533 bp in length
51029 51128: gap of unknown length
51129 52413: contig of 1285 bp in length
52414 52513: gap of unknown length
52514 54378: contig of 1865 bp in length
54379 54478: gap of unknown length
54479 55757: contig of 1279 bp in length
55758 55857: gap of unknown length
55858 57804: contig of 1947 bp in length
57805 57904: gap of unknown length
57905 59300: contig of 1396 bp in length
59301 59400: gap of unknown length
59401 60588: contig of 1188 bp in length
60589 60688: gap of unknown length
60689 62433: contig of 1745 bp in length
62434 62533: gap of unknown length
62534 64174: contig of 1641 bp in length
64175 64274: gap of unknown length
64275 65286: contig of 1012 bp in length
65287 65386: gap of unknown length
65387 67298: contig of 1912 bp in length
67299 67398: gap of unknown length
67399 68966: contig of 1568 bp in length
68967 69066: gap of unknown length
69067 70250: contig of 1184 bp in length
70251 70350: gap of unknown length
70351 71676: contig of 1326 bp in length
71677 71776: gap of unknown length
71777 72899: contig of 1123 bp in length
72900 72999: gap of unknown length
72999 74690: contig of 1691 bp in length
74691 74790: gap of unknown length
74791 75850: contig of 1060 bp in length
75851 75950: gap of unknown length
75951 77055: contig of 1105 bp in length
77056 77155: gap of unknown length
77156 78265: contig of 1110 bp in length
78266 78365: gap of unknown length
78366 79576: contig of 1211 bp in length
79577 79676: gap of unknown length
79677 81067: contig of 1391 bp in length
81068 81167: gap of unknown length
81168 82201: contig of 1034 bp in length
82202 82301: gap of unknown length
82302 83371: contig of 1070 bp in length
83372 83471: gap of unknown length
83472 84840: contig of 1369 bp in length
84841 84940: gap of unknown length
84941 86137: contig of 1197 bp in length
86138 86237: gap of unknown length
86238 87452: contig of 1215 bp in length
87453 87552: gap of unknown length
87553 88718: contig of 1166 bp in length.

FEATURES
source
1..88718
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-68E7"

BASE COUNT 21517 a 20394 c 20073 g 22076 t 4658 others
ORIGIN

Query Match 15.08; Score 446; DB 2; Length 88718;
Best Local Similarity 78.0%; Pred. No. le-107;
Matches 552; Conservative 0; Mismatches 150; Indels 6; Gaps 1;

OY 2077 AGCTTAAGTTAAGTTAAGCGGCGTACCTTCCCGACCAACGCTCGAGTTGT 2136

DB	Accession	Definition	Version	Keywords	Source	Organism
Db 144126	GAAGGTACACCCCTGCTGCTGCGGGAGCCTGCTGCTCACACGCGACGTTCAGAAAGCATGGA					144185
Db 2671	CCCAACAAAGGCGCCAGTGTCCACACATGTGCGCAAGCCGACAGACAGCAGCAGTGTGCTT					2730
Db 144186	CCCAACAAAGGCGCCAGTGTCCACACATGTGCGCAAGCCGACAGACAGCAGCAGTGTGCTT					144245
Db 2731	TTCCAGTGGGTGATGAGAACACCCCTCCAGG					2762
Db 144246	TTTCAGTGGGTGATGAGAAATGTGGCCACAGG					144277
RESULT 7	AC069534	213913 bp DNA linear	1	HTG 16-OCT-2001		
AC069534/C						
LOCUS						
DEFINITION	Mus musculus chromosome 11 clone RP23-520L, WORKING DRAFT SEQUENCE.					
ACCESSION	AC069534					
VERSION	AC069534.20	GI:14547767				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILLTOP.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 213913)					
	Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C., Dederich, D., Thomas, S., Okunou, G., Carllock, C., Garner, T., Addison, S., Pace, A., Williams, G., Bonini, D., Brooks, A., Brown, J., Bunay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z., Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S., Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R., Gorrell, J., H., Gunaratne, P., Haller, G., Hernandez, J., Hognes, M., Hosack, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S., Kovar, C., Liu, J., Liu, W., Louised, H., Lozano, R., J., Martin, R., Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S., Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogub, M., Parish, B., Perez, L., Relter, D., Say, J., Shen, H., Vasquez, L., Watlington, S., Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A., Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstein, G., Worley, K. and Gibbs, R.					
TITLE	Unpublished					
JOURNAL	Direct Submission					
REFERENCE	2 (bases 1 to 213913)					
AUTHORS	Worley, K.C.					
TITLE	Direct Submission					
JOURNAL	Submitted (02-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA					
COMMENT	On Jun 25, 2001 this sequence version replaced gi:13162471.					
	----- Genome Center -----					
	Center: Baylor College of Medicine					
	Center code: BCM					
	Web site: http://www.hgsc.bcm.tmc.edu/					
	Contact: hgsc-help@bcm.tmc.edu					
	----- Project Information -----					
	Center project name: MAFN					
	Center clone name: RP23-520L					
	----- Summary Statistics -----					
	Sequencing vector: M13; 108821					
	Chemistry: Dye-Primer Bodypy: 49% of reads					
	Assembly program: Phrap; version 0.990329					
	Consensus quality: 215914 bases at least 440					
	Consensus quality: 218176 bases at least 430					
	Consensus quality: 219553 bases at least 420					
	Estimated insert size: 215987; sum-of-contigs estimation					
	Quality coverage: 8.2x in Q20 bases; agarose-gel estimation					
	Quality coverage: 8.2x in Q20 bases; sum-of-contigs estimation					

	* NOTE: Estimated insert size may differ from sequence length					
	* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).					
	* NOTE: This is a 'working draft' sequence. It currently					
	* consists of 5 contigs. The true order of the pieces					

	* is not known and their order in this sequence record is
	* arbitrary. Gaps between the contigs are represented as
	* runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence
	* as soon as it is available and the accession number will
	* be preserved.
	1 80577: contig of 80577 bp in length
	*
	* 80578 80677: gap of unknown length
	* 80678 157388: contig of 76711 bp in length
	* 157389 157488: gap of unknown length
	* 157489 192191: contig of 34703 bp in length
	* 192192 192291: gap of unknown length
	* 192292 206716: contig of 14425 bp in length
	* 206717 206816: gap of unknown length
	* 206817 213913: contig of 7097 bp in length.
FEATURES	
Source	location/Qualifiers
	1..213913
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/chromosome="11"
	/clone="RPZ3-5201"
BASE COUNT	58068 a 51330 c 50253 g 53853 t 409 others
ORIGIN	
Query Match	14.6%; Score 434.8; DB 2; Length 213913;
Best Local Similarity	77.9%; Pred. No. 1.1e-104;
Matches 539; Conservative	0; Mismatches 147; Indels 6; Gaps 1;
OY 2077	AGGTAAAGTTTAAATTTAAGCGGGTGACCTTCGCCGACCATTGCCTGAGTTTGT 2136
Db 88393	AGGTGAAGTTGAATTTTAAGTGGTGATGACCTTCCCAAGCATGCCCCCTGGAGTTTGTc 88334
OY 2137	TGCTGCATCGCGGGATGCGAGCACACCTCGAGGAGATCCTTGACTGAGATTTTCAGGG 2196
Db 88333	GCGTCATATAGGTGATGTGATGATGAGACTTTGAAGAATACTTGCGCTCGGATTTCCACGG 88274
OY 2197	GGCCCCCCCAGGGCTTAGCCAGCCCTCTGCGCGCCCTGACAGCTAACCACTCCCTGAACAG 2256
Db 88273	GCTCCCCAAGAGCTAGGCCAGCCCTCTGCGCACCTGCAAGGCCAGCCAGCCCTCAACAG 88214
OY 2257	ATGGACACAGCAGCA-----CCCCAGCCTGCTGACAGACAGACAGCTGGTCTCA 2310
Db 88213	ATGGACACAGCTAGCACAATACCTGTCTCACTCTGTGTTAACAGACATATCTAGACTTCA 88154
OY 2311	AAGGCTGTGGCCCCAGACCCTCCACACCACCGCTGCTGTGAAGAACAATTCTGTGACC 2370
Db 88153	AAGACAGTGGCCCAAGCCTTGCTACACCCACTACTGCTGTGAAGAACAATCTGTGACT 88094
OY 2371	TGCACATGTGGCCAGGAGGCTGTGCTGCTCACTGTCGTAAGGAGAGGGCCCCAACCGGGGC 2430
Db 88093	TGCATTTGTGGCCGGAGAGCTGTGTCTCTCACTGTGTCCCAAGCAGAGGCCCCCAACGAGGC 88034
OY 2431	CGGAGTCTTTAAGTGTCAACGGAGGTAGTCACAATTCTTCTGTGGGCAGACAGCCC 2490
Db 88033	CGGCATTTCTCAAGATGTCAGTAATATGTTGACHTGCACTTCTTCTGTGGGCTACAGCAGC 87974
OY 2491	AATCGGGAGACAGAGGGGCTCTGTGCTTGTGGCATATACACCCCTGTGGGCGCTCTCCGTGGGA 2550
Db 87973	CATTCCACTGAGAGGGGAGCCCACTCTCATATGAGAGCCCAAGGCAAGCTCTGTATAGA 87914
OY 2551	TGCCACACAGGCCCAAGGATCACCTTAGTGGGTTTGGCAACCCGTGATGAGCAAGTGGT 2610
Db 87913	TGCCCATCTCATATGATGAGAGCACAATGATGATTTGGCAGACCTTGGGACGACAGTAT 87854
OY 2611	AGTGGACATCTCTGCTTGTGACGCCAGCCCTCCGCTCACAGGACTGTGCAGAGATGGA 2670
Db 87853	GGAAGTACACCTCTGCTGTGGGAGCTCTGTCAACAGGACTGTTCAGAGGATGGA 87794
OY 2671	CCCAACAAGGGGGCGAATTCACACATGTGCGCAAGCCGAGAGAGACACAGTGTGGCTT 2730
Db 87793	CCCAACAAGGACGCAATTCACACCTGTGTGCAAGCCACGAGACAGATGTGGCTTC 87734
OY 2731	TTCCAGTGGGTGATGAGAACCCGCTTCAG 2762

Db	29701	TTTTCCCTAAAGACCTTGCAGGGAAACGAGGCGCAGTGTACGAACCTGGTCTCCGGCACT	29642
QY	1256	TCCTGGCTTGGCTGCTCCAGGATGCTCAGAGGCGCAGSAGACCACTGGAGATCGACATCG	1315
Db	29641	TTCTTCGCCCTGGGTGACGAAAGATGCGATCGGGCTCCGAAACCCCTTGCTCATATCGACATCG	29582
QY	1316	CTCAGGAACGGCTTTGTGGACCCATGGCCATCATGATTTCTGGGCCGGAACATATCTGSAATGTT	1375
Db	29581	CCGGGAGAAAGTTACACGGCCCAATGGTCTTGCTTTATTCACGACGAAACTACCTGATGTTT	29522
QY	1376	ATCCATATGATCACTAGCGAGTGCACAGATCTCCCTGTCTATGAGCAGAGATCCCATTTTC	1435
Db	29521	ATGTTTATGACAAAGTGGAGTGGCAGCAGATCCATCACTATGAAAACGGGACGCGTTTG	29462
QY	1436	AGCCGACGACCGGTGGAGATGTGGTGCAGCGGGAGACCAAGCCCAAGCTGCTACCGSAG	1495
Db	29461	AGCCGACGAGAGTGTCTGCTCCACGAAAGGTGGCCACCACTGCCCCGCGCTGACGACGAAG	29402
QY	1496	CCGACCTCATTTGCCCTCCTATGAGAGAAAGCATGGATGGATGTGATAGCATATCCACTATGGGGAGC	1555
Db	29401	CGGATTTGATTGGTCACTGATGTGAGAAAGCATGGGATTTGGTATCAAGATCCACGACGCTGAGAC	29342
QY	1556	ACATCGAGACCATCAAAACCCGAGATGATCGTGGGGCTCAACCCACAGACAAGCGCTTCCTCC	1615
Db	29341	ACATTAACACATCATCAGSAGACCGCGATCATCTGGAGTCTGTGATTAAGGGCTTTTGGTAC	29282
QY	1616	CTGGGCAACCTGGGCACTGGGACTTGTGSAAGGTTATGATTTCCATGGGCTATGAAATGTCTA	1675
Db	29281	CTGGAGTATTGGATGGAGGACTGTATGAGGGGATACGATGCCATGGAAGCTGGCCCTTGCCA	29222
QY	1676	AGCCTGACCTCCGGGCTCAACTGCAAGCTGATCTTAAGCTCATCTGTGTATGSCAAAAG	1735
Db	29221	AGCCACAACCTTCGTGCTGATGTTGTAATTTGAACTGTGAAACTAATTTGCCAGGGGGCAGAAAG	29162
QY	1736	ACAATTTGTGTTCTAAGGACGACGACAGTGCAGAAATACAGGAGG	1780
Db	29161	ATCCCAAGGTGTCTCTGACCGACACGATGATCCCAAGTATACAGGAGG	29117

LOCUS	157766 bp	DNA	linear	INV 08-NOV-2003
DEFINITION	Drosophila melanogaster, chromosome 2L, region 37D-37E, BAC clone BACR27M12, complete sequence.			
ACCESSION	AC009017			
VERSION	AC009017.1			
KEYWORDS	GI:16798949			
SOURCE	HMG.			
ORGANISM	fruit fly.			
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyridroidea; Drosophilidae; Drosophila.			
AUTHORS	1 (bases 1 to 157766) Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Banazon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champe, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Fartan, D., Ferrieres, S., Frishe, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, C., Houch, J., Hoskins, R.A., Hostin, D., Howland, T.J., Idegawa, C., Jatali, M., Kruse, D., Li, P., Mettel, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nuno, J., Pacle, T.V., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouaneavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svitskas, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.			
TITLE	Sequencing of Drosophila chromosome 2L, region 37D-37E			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 157766) Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Banazon, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champe, M., Davenport, L.B., Dietz, S.M.,			

TITLE	JOURNAL	COMMENT	FEATURES
Query Match	Best local similarity	Matches	14.6%; Score 433.4; DB 3; Length 157766; 55.7%; Pred. No. 2.66-104; 0; Mismatches 671; Indels 119; Gaps 4;
BASE COUNT	44324 a	35420 c	3514 g 42508 t
ORIGIN	<p>1. 157766</p> <p>/organism="Drosophila melanogaster"</p> <p>/strain="y: cn bw sp"</p> <p>/db_xref="taxon:7227"</p> <p>/chromosome="2L"</p> <p>/map="37D-37E"</p> <p>/clone="BACR27M12 (D1301)"</p> <p>/clone_1lb="RPCI-98 (Rosewell Park Cancer Institute</p> <p>Drosophila melanogaster BAC library, partial EcoRI in</p> <p>PRC63.6)"</p> <p>location/Qualifiers</p>		
112	CGAGAGAGAGACCTTCAAAATTCACACAGATCTATGATATTTGATATCATCTGTATGGC	171	11
85416	CAGCGCGAGAGCTACTCGTGTACACAGAGCTTCGATTTGAGAGCCCTGTGGCGCGT	85475	11
172	CAGAGTGTACACGTGATGATGACTTCATTTCTGTGACATTTACTAGCGCTCATGATTTCCAG	231	11
85476	CAGATGCGCAAAATGATTATATACCTCGTATACCGGCCACATGATGCAGCTGGCATTTCCAG	85535	11
232	ATGCGAGTTTCAGAAATGCGACAGCTGCACACCCCTCTTCTCTTTGAGACGAAATTTGAA	291	11
85536	GTGCTTTACAAAATTTGGAGAACGTTGATCGCGGCTCTTTATTGACAGCGCCGGTGAAA	85595	11
292	AAGTACTGCCACAGAGATTTTGTAGACATCAAGAAAATTGGAAACGAGACACTCGCCAG	351	11
85596	AAGGCGCTGGGCTGAGATTATGAGCGCTATCAAAAAGACCTTGGAAACGCGAGGTGGCGCGC	85655	11
352	TTCGACGCTGCGGATGCTGACATGACATGATATGAGAAAGCGAAAACATGCGGGTTTGA	411	11
85656	TGTGAGGGATGATTATCTGAGCGGATTCGATCGCGGCGCGAGAACATTTGGCTACAG	85715	11
412	ATTATTCACAGTGTGTAAAGGCTGTAAAGCCCAATCTGCAGGTGTTGCGAGCCGATTTCT	471	11
85716	ATTATTTGATGTGTGTCGGCGCTATCAACCGAATTTTTCGGTTTATCGTGCACATTTTCG	85775	11
472	GAGATCACACCCATGCCGCTGAGACAGCTTGTGAAAACCTGCACCGAGCTGTATCAGG	531	11
85776	GAGATTTACACGCGTGGCGGTCTCTCGGCTCTTGCAGCATTTG888GACGCCGACAAAAG	85835	11
532	GTGAGCATGCTGTGGATGTTGAGGACGAGCTGCAGCCGAGAGATTGGAGCTTTCCT	591	11
85836	CAGAGCATGAGTGTGATGTGCGACGGAACCTGGATTTGCGACATGGGTGGCGCATCC	85895	11

[illegible][illegible]

RESULT	10
AE003663	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

LOCUS	299449 bp	DNA	linear	INV 04-OCT-2008	
DEFINITION	AE003663	Drosophila melanogaster genomic scaffold 142000013386055	section 56		
ACCESSION	AE003663	AE002690			
VERSION	AE003663.2	GI:10728874			
KEYWORDS	fruit fly.				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 299449)				
AUTHORS	Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Blandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,J.H., Blazer,V., Blasey,C., Champ,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor,Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J., Andrews-Plamkoc,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Bernan,B.P., Bhandari,D., Bolshakov,S., Borrova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brotlier,P., Burtils,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahle,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durin,K.J., Evangelista,C.C., Ferrara,C., Ferrieri,S., Fleischman,W., Foster,C., Gabrielian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibbegan,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Matel,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Mekuklov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nussken,D.R., Pauley,J.M., Palazolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reiner,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siding-Klamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskis,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wasserman,D.A., Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q., Ye,J., Yen,R.F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,W., Zhou,X., Zhu,X., Zhu,S., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.				
TITLE	The genome sequence of Drosophila melanogaster				
JOURNAL	287 (5461), 2185-2195 (2000)				
MEDLINE	20196006				
REFERENCE	2 (bases 1 to 299449)				
AUTHORS	Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.				

TITLE Direct Submission
JOURNAL Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT On Oct 9, 2000 this sequence version replaced gi:7298582.
FEATURES
source location/Qualifiers
1..299449
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2L"
join(619..833,3530..3863,3923..4604,4675..5176,5254..5430,
5487..5895)
/gene="Sd"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT28175"
/db_xref="FLYBASE:FBan0009999"
/db_xref="FLYBASE:FBgn0003346"
<619..>5895
/gene="Sd"
/note="CG9999"
/map="37E2-37E3"
/db_xref="FLYBASE:FBan0009999"
/db_xref="FLYBASE:FBgn0003346"
complement(join(1516..2125,2209..2429,2490..2628,
2687..3005,3071..3339))
/gene="Hs2st"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/product="CT28767"
/db_xref="FLYBASE:FBan0010234"
/db_xref="FLYBASE:FBgn0024230"
complement(<1516..>3339)
/gene="Hs2st"
/note="CG10234"
/map="37E3-37E3"
/db_xref="FLYBASE:FBan0010234"
/db_xref="FLYBASE:FBgn0024230"
complement(join(1882..2125,2209..2429,2490..2628,
2687..3005,3071..3197))
/gene="Hs2st"
/note="Hs2st gene product; Nucleotide sequence of the
Celera sequence differs from the published sequence for
this transcript"
/codon_start=-1
/db_xref="FLYBASE:FBan0010234"
/db_xref="FLYBASE:FBgn0024230"
/protein_id="AAf53800.1"
/db_xref="GI:7298583"
/translation="MFRKLKMWLLRPHTWLLIACAVTCAGYWLMSRIRLHAF
KPLSGDLSLSPQHSSTDDPDEFEHLVLYLNRPVKTGSTSVNATYDLCKRNKH
VLIHNTANHHVLSLPNQIOFVRNVSRWHEMKPALYHGHAFIDFSKFOIAHKFTYIN
LVRRPLDLRVSYTYFLFGDNYRPNLKRAGNKTITFDECVOVKOPDCDPRNMNLIQIP
FPCGHAACEWPEGSWALDQAKRNLYEVFLVGTQMEYFVLDLERSLRIIFGFEF
HYHNSKSHLRVTSKRLPSESTIKSIQTKIKQWENDLYDFALAQEFPMKKKLMQD
NKHVKPMYKRIPIPK"
join(3764..3863,3923..4604,4675..5176,5254..5430,
5487..5816)
/gene="Sd"
/note="Sd gene product; Nucleotide sequence of the Celera
sequence differs from the published sequence for this
transcript"
/codon_start=-1
/db_xref="FLYBASE:FBan0009999"
/db_xref="FLYBASE:FBgn0003346"
/protein_id="AAf53801.1"
/db_xref="GI:7298584"
/translation="MSTFNFASMAAOLGQEGISFENKVLSTMTAADVDVVALNKO
TVHYLMDGNTLGEBAKATIGSLKHPFKRALMKNMTGRLISIFPALHGLGA
LIVAAGLTVLSDNALGPNMKGLELLRSPVCSYLOELLCLNCGLGEGSMISR
ALIDHANANKAGFPLOLRFVIGSRNRLDAGATEMATPAFLTKREMSFDCILKTIY
EGVEALAESFKNHPHLVLMNNDNTLSSEGAETIAELPILPLIREMSFDCILKTIY
AYHGEALGSRNLELYIDLGPNINDGLVLYVMNGNPKLRIINLNDGNSGEEBS
EKIISMSKRIPTAAALQIPQHOEEEDLEDEYQADKQADYEEEEEVHEHNDTTEAD

EDSEGDDEDEDEDEEYSNAEETAVYTTNAYTKLFENDTNSMASETPAVANKTIS
QKCPKFKCLSLQKPCSQDEFDLMDNKLNLQISIVNDPTDNHLLLVFTTLCAHL
SQSKPAALDLDAKLYQATFDYAIKTKQETRYLVNLMQRLRLPKKEVHSDYVKNCR
FALREALKOPTFADNKNISKFTLEGAES"
complement(join(6316..6639,6721..9758,9825..10002,
10075..10248,10306..10863,11829..>12084))
/gene="Top2"
/product="CT28703"
/db_xref="FLYBASE:FBan0010223"
/db_xref="FLYBASE:FBgn0003732"
complement(<6316..>12084)
/gene="Top2"
/note="CG10223"
/map="37E3-37E4"
/db_xref="FLYBASE:FBan0010223"
/db_xref="FLYBASE:FBgn0003732"
complement(join(6316..6639,6721..9758,9825..10002,
10075..10248,10306..10863))
/gene="Top2"
/note="Top2 gene product"
/codon_start=-1
/db_xref="FLYBASE:FBan0010223"
/db_xref="FLYBASE:FBgn0003732"
/protein_id="AAf53802.1"
/db_xref="GI:7298585"
/translation="MENGKALSTEDQYOKKSQLEHLLRPDSITGSVEFTKELMNY
DNSONRMVOKESISFVPGLYKTFEDILVNAADNKKORDKSMNTIKIDIPERMVSVNN
GOGIPVIMHKEOKMYVPTMIFGHLLTSDYDKRTYGGNKGAKCNFTSFVLY
ETATREYKSKSFQWNGNMKGASDVQALIKDFNGTGYRTITFSPDLAKFMDLSDIYA
LMSRRAYDVASASKGVSYFLNGKLGIVHEVARNRMVACQPSRGGQOYFVNSIVT
KSGRRYHDVYDNLIKOLEVLRKNNKNGGINTIKRPQVRNHHVYVKNLLEPTDSOT
KENNTLOOKGFSGKCTLSEKFINNMSGSIVSVLWAKMFAQNDIKATGKRSKIK
GIPKLEDANEAGKNSIKCTLILLEGGSASLAVSGVIGRDLGVFPLGKLLNR
EANKQLESENAINNLCIKIGLOYKKYILJEDDKTURYGVKMTDODDGSHTKGL
LINFITHNMPELRLPFLERITPITYAKTKNEBLSFSLPEFEKMNDAHNTYNI
KYVGLGSTSKAKKEYFODMDRHRILFKYGVSDVDSIYVAFSKHIESKVVLTJH
MDPKRRRELGLPRVLYTKGTGKITADPTINELVPSNADNRSRISLYDGLKPO
RKVMTFTPKRNDKREYVAVQSSVVAEMSAVHHEVSLQMTIVYLAOPFGANNIL
EPKQFSTRLSGSKDCASARYIFETIMSPLERLVHPDLDDLQVDDGQITEPLNLT
PIIPMLVLCNGAIGIGTGMSTKISYNNPREIKNLNRMINGEPVMPHWKKNFLGRHE
YVSDGRYIQTGNIQILISGNRLSELVGVVTOYKKNVLEPLISNGTEKYVGLISEYR
EYHDTTVYRVFVISAPEFERITHAEGGVYFVLTTLTSTNOMHAFDONNCLRPPT
AIDILKEVYKLRREYVAPRDFIVQGLTAODRILSDARFLTECKEIKIYENKORAK
MCDLLKRGYRPDPYKEMORITKMEDEQADDEDEDEEAPSVSSAKKKEVDPK
AFKRLTVYKRFDYLLGSMWMLTEKKNELLKQDYTLUSELSLRKTPKMLVDDLD
ALBSKLINEVEKERAEEOGINLKTAKLKQKSSASAGRKYSKSGGAGAGDVPDD
GEPEYFKITTEELIKKMAAAKVAQAAKPEKPKPKPKPKPKPKPKPKPKPKPKPK
EVDEFDLMEGSKTSPRAKAVYKKEPEKKPKPKENGGLKOSKIDFSKAKAKKS
DDVYEYTPRAERGRQASKIDYSSLFSPDEEDGNGVSDDGNASDDSPKRPKAK
RGRDESSGGAKKAPPKKRAVYESDDDLIEDDDDDSDDFNC"
join(13443..13558,14987..15146,15208..15302,15402..15746,
15806..15984,16054..>16171)
/gene="CG10026"
/product="CT28189"
/db_xref="FLYBASE:FBan0010026"
/db_xref="FLYBASE:FBgn0032785"
<13443..>16171
/gene="CG10026"
/map="37E3-37E4"
/db_xref="FLYBASE:FBan0010026"
/db_xref="FLYBASE:FBgn0032785"
join((15011..15146,15208..15302,15402..15746,15806..15984,
16054..16171)
/note="CG10026"
/gene="CG10026 gene product"
/codon_start=-1
/db_xref="FLYBASE:FBan0010026"
/db_xref="FLYBASE:FBgn0032785"
/protein_id="AAf53803.1"
/db_xref="GI:7298586"
/translation="MPTIEKLNITTEEVPEHILRLAQEGECPSSKDKVITEOPRNYI
LEHNEDCPHNSDAKILYKFLARARWKILENSKLLCSYRRREQOKSYEVNRPDLRHH

Query Match 14.6% Score 433.4; DB 3; Length 299449;
 Best Local Similarity 55.7% Pred. No. 2.8e-104;
 Matches 995; Conservative 0; Mismatches 671; Indels 119; Gaps 4;

QY 112 CGAGAGAGAGACTTTCAAATTCACAAAGATCTATGATTAATTCATCTGTATGCG 171
 DB 57035 CAGCGCGAAGGCTACCGGTGTCACAAAGTCTTGATTTGAGAGCCCTGTCGCGGT 57094

QY 172 CAGATGTTACCATGCTATGATTCAGTTCTGACATTTACTGGCTATGATTTCCAG 231
 DB 57095 CAGATGCTCAAAATGCTTATGATTCCTGATCCGGCCACATGATGAGCTGGATTCAG 57154

QY 232 ATGCAATTTGAAATGCGAGAGCTCAACCTCTGCTCTTTAGAGAGAAATTTGA 291
 DB 57155 GTGTCTTACAAAATTTGAGAAAGGTGATCCGCTCTTTATTCAGCGCGCGTGA 57214

QY 292 AAGTACTGCCACAGAAATTTTGTAGACATCAAGAAAATTTGAGACGAGACTCGCCAG 351
 DB 57215 AAGGCGTGGGCTCGGATTTATGAGCTTATCAAAAGACCTTGAACGCGAGGTGCGCGG 57274

QY 352 TGCAGAGCTCTGTGATCTGAGACTGATGATAGAGAAAGCGAAAACATCGGCTTTGAG 411
 DB 57275 TGTCAAGGATTTGATTTATCTGAGAGATTTGATGCGAGGCGAGAAACATTTGCTACGAG 57334

QY 412 ATTTATCCAGCTGTGTAAGGCTGTAAAGCCCATCTGAGGCTTTGGAGCGCGATTCCT 471
 DB 57335 ATTTATGATGTGTGTGCTGCGCTATCAACCGAATATTTGCTTTATGCTGCTACTTCTG 57394

QY 472 GAGATCACACCCCATATGCTGAGAGAGCTGTGAAACCTGACCGCTGATCAGAG 531
 DB 57395 GAGATTACACAGGTGTGCGGCTGCGGCTCTGACACAAATTTGGGCGACCGCGAAG 57454

QY 532 GTGAGCGATGCTGTGATGTGAGGACAGAGCTGACCTGAGATTGAGAGCTGCTTACT 551
 DB 57455 CAGAGGATGATGATGTGCTGAGAGCTGAGAACTTTGGAATTTGGAATGCTGCGCATCACC 57514

QY 592 AGGTTCCAGACCTGTGCGGCTTCAGAGATTTTCTGAGGTGCTGGACAGAGAGCTATC 651
 DB 57515 CGGTTTCAGACATGCTGCGCTACAGCTGTCTCCCGAAAAGTATGCTGATTAAGCTCAT 57574

QY 652 AGTTAGGAGAGCTGCGAGTTCCTCCACAGCTGGGCTTGTGTGAGGCGTTCAAGCCAT 711
 DB 57575 TCTATGAGAGCTGCCAAATTTCCACCCCTGGATGCTGTCGCGAAGCGGTACAGAGAT 57634

QY 712 CAGGCTTTTGTACCAAAATCTTCCACAGAAATTA----- 746
 DB 57635 GAGGCTTCTGATTCGAGGCAATTTTGAAGATTAAAGGTAGAGAGAGATGACATACGA 57694

QY 747 -----AGTACTCATGACCAAGAGATGATGATGAGA 779
 DB 57695 TTGGAAAGCTATTAATTTGATTTCTAGTTCTGACACAAATGATGATGACGCTGGA 57754

QY 780 ATTCACTGGAAGAGAGCTCTTTTAACACAGCGCTTGGCTTCTATCATGTT 839
 DB 57755 GTTTAATTTGGGCGCGAAGACCGACTTTCGATTAAGAACTTGTGAATAATTTGCTCT 57814

QY 840 GTGTGTGAGAGATCCCATGCAACTGTGTAGAG-----GTCAAGTCTAAGCCCAAGAG 893
 DB 57815 GTGCTGTGCGGAGAACCGGATCTAGACGCTGTGAGAGAGCTGACCGTTAAGGCCAACA 57874

QY 894 CAGTGTGCGGCTCAAGCTTTGAGACTGTGAGACTTTGAAACCTGCTGCTGAAAGTT 953
 DB 57875 CAAATGCGCGCGCAACTTCGCTGATACCGTGAATGAGAAAGCTTGGCTCGAGAAACT 57934

QY 954 GAGATTAATTTGTAAGAAACCATGAGATTTGCTGAGAACTCTACATCAAGGGATCAT 1013
 DB 57935 CAACTATCTGCCAAGAGAGCATGACAAATAGCCGAAATAATTTATACCAAGGTTTCAT 57994

QY 1014 CAGCTATCCCGAAGAAACAATTTTCCACAGACTTAAACCTGACGCTTTGGT 1073
 DB 57995 CAGCTATCCCGTAAGAAACAACAGTTCTCTCAAGGATTTGCTGTGACACACTAGT 58054

QY 1074 GGAACGACAGACCCCGGATCCAGCTGTGGGGGCTTTGCCAAGAGATTCTAGAGCGGG 1133

DB 58055 TGAGATGCAAAAGGATTCACCGGCTGAGAGATTTGCCACGCGGTGATGTAGTGG-- 58112

QY 1134 TGTCTCCACCCCAACGAGTGAAGCAAGTCTGACCAAGCTACCTCCATTCACCCAC 1193
 DB 58113 -GGACCAATCTTCGAAGCGTAAACAAATGATGATGAGGCGCCATCTCCAAATTCATCCAC 58171

QY 1194 CAAA----- 1197

DB 58172 CAAGCTGGCAGAAAGTAAAGGAAATCTTTATTTAATCTAAACCGAACAATACATACCC 58231

QY 1198 --TACACCAACAATTTACAGAGATGATGACAGAGCTGTACAGATTATTTGCTGCGCAT 1255
 DB 58232 TTTCTCTTAAGACCTTCAGAGGAAAGAGCGGAGGTGATGCAAGTGTGCTCGGACT 58291

QY 1256 TCTGTGCTGTCTCTCCAGAGATGCTCAGGGGACAGACACAGTGAAGATGACATCG 1315
 DB 58292 TCTGTGCTGTCTCTCAGAGATGCTCAGGGGACAGACACAGTGAAGATGACATCG 58351

QY 1316 CTCAGAGAGCTTTGTGGCCATGCGCTCATGATTTCTGCGCCCAACTATCTGATGTGT 1375
 DB 58352 CCGGTGAGAAATTCACGCGCAATGCTGTGTTATTCAGCAAGCAAACTACCTAGATGTT 58411

QY 1376 ATCCATATGATCACTGAGTGAAGATGATCCCTGCTATGAGCAAGATCCCATTTG 1435
 DB 58412 ATGTTTATGACAAAGTGAAGTGTCCAGACAGATCCATCTATGAAACGGCAGCGTTTG 58471

QY 1436 AGCCACAGACCGGTGAGATGTTGACGCGGAGACACACCCCAACCTGCTACACGAG 1495
 DB 58472 AGCCACAGGAAAGTGTGCTCCAGAGAGTGTCCACACACTGCGCCCGCTGACGGAAG 58531

QY 1496 CCGACCTATGCTGCTCATGAGAGATGATGATTTGATGATGATGATGATGATGATGAT 1555
 DB 58532 CGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 58591

QY 1556 ACATGAGACATCAAGAGCGCGATGATGATGATGATGATGATGATGATGATGATGAT 1615
 DB 58592 ACATTAAGACGATCAAGAGAGCGGATGATGATGATGATGATGATGATGATGATGAT 58651

QY 1616 CTGGGACCTGCGGCTGAGAGCTTGTGAGAGTATGATGATGATGATGATGATGATGAT 1675
 DB 58652 CTGAGATTTATGAGATGAGAGCTGATGAGAGTATGATGATGATGATGATGATGAT 58711

QY 1676 AGCTGACCTCGCGGCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1735
 DB 58712 AGCCACAACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 58771

QY 1736 ACAAAATTTGTTCTTAAGGACAGCAAGTGCAGAAATACAGACAG 1780
 DB 58772 ATCCCAAGGTGTGCTGACCGAACAATGATACCAAGTACAGACAG 58816

RESULT 11
 AC018257
 LOCUS 73907 bp DNA linear HTG 09-DEC-1999
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
 pieces.
 ACCESSION AC018257
 VERSION AC018257.1 GI:6552934
 KEYWORDS HTG; HTGS_PHASE2.
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 73907)
 AUTHORS Adams, M. and Venter, J.C.
 TITLE Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 JOURNAL
 COMMENT
 This sequence was identified as CDM:10214069 by the submitter.
 For more information on this record e-mail to fly@celera.com.
 * NOTE: This is a 'working draft' sequence.

REFERENCE 1 (bases 1 to 2422)
AUTHORS Kim,Y.C., Lee,J. and Koo,H.S.
TITLE Functional characterization of *Caenorhabditis elegans* DNA
JOURNAL Nucleic Acids Res. 28 (9), 2012-2017 (2000)
MEDLINE 20221600
REFERENCE 2 (bases 1 to 2422)
AUTHORS Kim,Y.-C. and Koo,H.-S.
TITLE Direct Submission
JOURNAL Submitted (01-APR-1998) Biochemistry, College of Science, Yonsei
University, 134 Sinchondong, Seodaemunku, Seoul 120-749, Republic
of Korea

FEATURES
source Location/Qualifiers
1..2422
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/chromosome="3"
1..2422
/gene="TOP3"
13..2292
/gene="TOP3"
/product="DNA topoisomerase III"
/protein_id="AAC13567.1"
/db_xref="GI:3047013"
/translation="MKRALFVAEKNDVANGVAAILSTGANRREGSKENKIYTLNTE
LFGQOAI SVTSVGHMMNFQPHENMSNMOTASVRELFRAPVHVVTPEMKLIEQTLR
EOKORHDI LVTWDCDREGAIEAIKVCRODNRLDFRAFSEITKAITRAARN
LIRLDEKTVAAVDCRSELDLRIGSAFTROTJLHNRPRDLGONDSOTVSYSGOF
PRLGFTYDRKMIENFVSEPEFKLIYEHTRSHKVFELMDRNLPRDRTDILHDECK
ETKEAHEVYAKKPKSKWRPQALDVELELGLSKLRMSKQTMQVAEKLSKGFISY
PRETKHPFAGLMLPLVQOQTSNIMGFANSEVNLGVNPRGRSDEAHPIHPLK
FTEKHOLQGDVMKVELVYVHFLACYSOAQGEETVNVNLGTEKPHASGLRIIDMY
LKYYVEKMGNRLPLTYTEGEFDEFLKIGDGTQAPDLTFADLISLMDKYGIGD
ATHAHEIKETKREYIGVRDGLIPSLGIALVDYDMDGFAMSKPDLRANLEIGIK
ETCDGRKQOEVDLQIKRYATFVSEKRIYGLSOSLQRYLDKNNOAGGREGPRGG
GGPRPGGGGGGGGPTGPPAPKPPAKPRPRPRKSI SPVAKGHDDPENDITVLTSE
VFGSMNPAPKPRAPRKSAPKREQEEVEVCCQPEPRAVTKVYQKRGPKRKE
YTCSLPYTSEKCNPKFMA"

BASE COUNT 749 a 469 c 620 g 584 t
ORIGIN

Query Match 14.28; Score 422; DB 3; Length 2422;
Best Local Similarity 53.68; Pred. No. 1.6e-101;
Matches 956; Conservative 0; Mismatches 805; Indels 21; Gaps 3;

32 TGGGAAAGTCCTGTGTGGCGGAAAAAAGACGAGGGGCCAAGGGATGCGGACCGTC 91
14 TGAAGAAGCCTTATTTGTGGCCGAGAAAGATGACGTGGCAAGGAGTCCCTCGATT 73
92 TGTCAACGGTGCATGAGCGAGAGAGACATTTCAAAATTCACAGAATCTATGAT 151
74 TATCAATATGGAACGGAATCGACGAGGGGCGGACGAAGTTAACAAAGTTTACAGT 133
152 TTGATTTATCATCTGTATGCGCAGATGTTACATGTAATGACTTCAAGTTTCTGACAT 211
134 TGAATACCGAGCTTTTGTGACGCAACCGCAATTTGCGTGACAAAGTTTTCGGACATA 193
212 TACTGGCATGATTTCCAGATGCGATTTGCAAAATGCGAGAGTGCAGACCTTTGCC 271
194 TGATGAATTTCCAAATTTTCACAGAAATATGTGCAATTTGGCAACGGCGTCGATGCGGT 253
272 TCTTTGAAGAGAAATGAAAGTACTGCCAGAAATTTTGTAGACATCAAGAAAATTT 331
254 TGTTCGAGCTCCGTTTCGCGATGTTGTACACAGAGATGAAGCTATCGACAGACAT 313
332 TGGACAGAGACTCGCCAGTGGCGGCTCTGGTATCTGGACTGACTGTGATAGAGAG 391
314 TTAAGAGAACAGCCCAACGACACATATCTGTAGTCTGGACAGATTTGCGAGAG 373
392 GCGAAACATCGGGTTGAGATTTATCCAGCTGTGTAAGGCTGTAAAGCCCAATCTGACG 451

374 GTGAGCAATTTGGCGCTGAAATTTCTTAAAGTTTTCACAGATAGACACCGCGCGTGATA 433
452 TGTTCGAGCCCGATTTCTGATGATACACCCCATGCGTCAGGACGTTGTGAAAC 511
434 TTTTCACAGCAGATCTCCGAATTCACAAAGACGCATACACAGCAGCTCGAAT 493
512 TGACCCGACCTGATACAGAGGTGAGCATGTGTGATGTGAGCAGAGAGCTGACCTGA 571
494 TCATCCGACTCGACGAAGACCGCTTCGCCGATGTGATGTGCAAGTAATTTGATCTTC 553
572 GGATTTGAGCTGCTTTACTAGTTGAGAGACCCGCGGCTTCAGAGATTTTTCCTGAG 631
554 GAATTTGATCCGCTTTTACAAAGACTACAGACGTTGCAATTTAAGAAATCGATTGACAT 613
632 TGCTGG-----CAGAGAGCTCATCTACGTCAGTACGCGAGCTGCCATTTCCACAC 679
614 TACTGGGCAAAATATGACACATCCCAAGTATGACTATGCTCTGTCATTTCCCAACAC 673
680 TGGGCTTTGTGTGGAGCGGTTCAAGCATTTACAGGCTTTTGTACAGAAATCTTCCACA 739
674 TCGGCTTTGTACAGATCGTTATATAATGATGAGAACTTTGTGTCGAAACCATTTCTGA 733
740 GAATTAAGTAACTATGACCAACAAAGATGATCGTAGAATTCGAAATTCGAAAGGCATC 799
734 AGCTGATCGTGAGACACACCCGGAGACCCCTTAAAGTCGAATTTCTATGGATCGAATC 793
800 GACTCTTTAAACACACGCGCTTCCCTAGTTCTCTATCAAGTTGTGTGGAGATCCATGG 859
794 GACTTTTCGATCGAGACACCTGATATATTTCTTATGATGATGTAAAGAACCAAGAG 853
860 CACTGTGTGTAGAGCTCAGATCTAAGCCCAAGACAAAGTGGCGGCTTCAAGCTTGACA 919
854 CTCATGTTGAAAAAGTGGCGAAGGCGAAGACCAAGTGAAGACCGAGCTTTGGATA 913
920 CTGAGAGCTTGAGAGAGTGGCTTCTGGAAGTTGAGATAATAGCTAAAGAACCATGA 979
914 CGGTTGAAGTGAAGAAAGCTGGGAATCTCGAAGCTTAGAATTCGGCTTACGACATGTC 973
980 GGAATGCTGAGAAAGCTCTACACTCAAGGATCATCAAGTATCCCGAAGCAAGAACCA 1039
974 AAGTGTGAGAAAGTGTATGATGCAAGGATTTATTTGCTATCCGAGGACGAGAGAA 1033
1040 TTTTCCCAAGACTTAACCTGACGGTGTGTGTGGAACACACAGACCCCATCCAGCT 1099
1034 AGTTCCGGCTGATTTAACTTGACACCGTGTGACAGCAGACACAAATCGAATATTT 1093
1100 GGGGGCCCTTGGCCAGAGCATTTCTAGAGCGGGGTCGCCACCCCAAGCAAGTGAACA 1159
1094 GGGAGATTTTGTATGAGGTTCTTACAAACGGT---TAAATCCGAGAAACGGACGA 1150
1160 AGTGTACCAAGCTCACCCCTCCATTACCCACCCCAATATACAC-----CAACAATTAC 1213
1151 AAACGAGAAAGCCCATCCACCGATTCATCCAGCAAAATTCAGGAAACACCAAGTTAC 1210
1214 AGGAGATGAACACGACTGTACGATTTATTTGTTGCGATTTCTCGCTGCTGCTCC 1273
1211 AAGGACAGATTTGAAAGTCTATGAGCTCGTTGTCGATTTTGGCGTGTGCTCC 1270
1274 AGGATGCTCAGGCGAGAGACCAAGTGAAGATGACATGCTGACAGAGCGTTTGTGG 1333
1271 AGGATGCTCAGGCGAGAGAAACAAATGTTAATTTTAAAGTTTGAAGCAGAAATTTCCATG 1330
1334 CCCATGCGCTCATATCTGTGCGCGGAAACTATCTGATGTATCATATGATCATCTGA 1393
1331 CTTACGCGCTTCGAATATGAGATATGAGATATTTGAAGTCTATGATGATGAGAAATGG 1390
1394 GTGACAGATCTCTCTGTCTATGAGCAAGATCCCACTTTCAGCCACAGCAGCTGAGGA 1453
1391 GTAATCGACTTTTCCCACTTACACGGAAGGAAAGCAATTCAGTAATTTGAGCTGAAGA 1450
1454 TGGTGAAGGAGGAGACGACCCCAAGCTGCTACCGAGGCGGACCTCATTCGCCCTCA 1513

Db	1451	TTGGAGATGGAACAAACACAGGCGCCGGATTTTTTTGACTAGCGGATTTAAATCTCACTAA	1510
OY	1514	TGGAGAAAGCANTGGCATTTGGTAGGATGCCACTCATGCGAGACACATCGAGACCATTCAAG	1573
Db	1511	TGGACAAATATACGGAATTTGGAAACCGAGCAACACACGCCGACAAATTTGGAAATATCAAA	1570
OY	1574	CCCGGATGTAGCTGGGGCCACCCGACAGAAAGGTTCTCCCTGGCGACCTGGGCATGG	1633
Db	1571	CACGAGATATATTGGAGTAGCTCCAGATGGAATTAATTCATCATTTCTCGGTCTGG	1630
OY	1634	GACTTGTGGAAGTTATGATTTCCATGCGCTATGAAATGTCTAAGCTCGACCTCGGGCTG	1693
Db	1631	CACTTGTGCGACGGTATGATGATATGCGGCTTTGCATGTGAAGCCCGGATTTACGAGCCA	1690
OY	1694	AACCTGGAAGCTATCTGGAAGCTGATCTGTGATGGCAAAAAGCAAAATTTGGTGTTCTAA	1753
Db	1691	ATTTGGAAATCGGCTGGAAGAATAATTTGGGACGGCTGCACACAGAAACAAAGATTTTGG	1750
OY	1754	GGCAGCAAGTCAGAAATATCAAGCAGGTTTTCATTGGAACGG	1795
Db	1751	ATGAGCAAAATCGCAAAATATCGAGCCATTTTTGTGAGAGTCTG	1792

[illegible]

	KSYPLCPYCYNHPPFRDMKKMGNGNECTHPSQHSLSMLGIGCVESGVLYLDP	
	GPKKXVAVCNKNVCNAHFENAHRYRVSADTCSVEEALLVDFFKASAPJCETOM	
	GCVEDPVDFPOLVELLKHAISCHPMHRGPGRQRGRBARBRPGRPNPRRKDMA	
	LAAFAVV"	
variation	2447.. 2448 /gene="TOP3D" /replacement="c"	
BASE COUNT	629 a 829 c 806 g 520 t	
ORIGIN		
Query Match	9.1%;	Score 270.6; DB 9; Length 2784;
Best Local Similarity	49.5%;	Pred. No. 5.3e-61;
Matches	882; Conservative	0; Mismatches 869; Indels 30; Gaps 6;
Oy	32 TCGGAATCCTCTGTGTGGCCAAAAAAGACGAGCGGCCAAGGGGATTCGCCACTGC	91
Dd	115 TCAAGCTGTCTCATGTGTGTGAAAAGCCCTCTTGACAGTCATTGCCAATAACC	174
Oy	92 TGTCAAACGGTCGCATGAGCGGAGAGAAGCACTTTCAAATTCACAAGATCTATGAT	151
Dd	175 TCCTTAGAGGGAGCGCTGTCTCACACAAAGGCTGAAGGGGCCGTCACTCCAAGGT	234
Oy	152 TTGANTATCATCTGTATGAGCCAGAAATGTTAACATGCTATACATCTTCTGGACATT	211
Dd	235 ACACGTGGACCTTCTCTGGCCAGCCAGTGCCTTCMAATGATACGCTGTCTGTGTACG	294
Oy	212 TACGTGGCTCATGATTTCCAGATGCGAGTTTTCGAAATGGCAGAGCTGCAACCTCTTGCC	271
Dd	295 TCATATACCTCGATTTCCTCGSAAATATCAACAAATGGAGCAAAAGTGCACCCGACAGC	354
Oy	272 TCTTTG---AAGCAGAAATTGAAAGTACTCGCCCAGAGAAATTTGTATACAAAGAAA	328
Dd	355 TGTTCAGCCAGCGCTCCACGAGGAAAGAAAGTAACCCAGCTGAACATGTGTGAAGT	414
Oy	329 CTTTGGAAACGAGAGACTGCCAGTCCAGGCTCTGTGTATCTGAGTACTGTATAGAG	388
Dd	415 TCCTCAGGTGAGGGCAGAGAGCTGCAGCTACATCTGTCTGTGCTGTGAGCGCAAGG	474
Oy	389 AAGGCAAAACATCGGCTTTGAGATTATCCAGTGTAAAGCGTAAAGCCCAATGTGC	448
Dd	475 AGGGGAGAACATCTGCTTTGAGTCTTTGATGCTGTCTCCGCTCATGAACAAAGCCCC	534
Oy	449 AG-----GTGTTCGAGCCCGATTTCTTGAGATCACACCCCATCCGCTCA	493
Dd	535 ATGTGTGGGAGAGAACCGTGTTCGGGGCCAGGTTAGTTCATCAGCGACAGACATCT	594
Oy	494 GCACAGCTTTGTGMAAACCTGACCGAGCTGTATCAGAGGGGTGAGGATGCTGTGATGTA	553
Dd	595 GTAAATGCCATGGCTCTTAAGCGAGCTGTGACCAACAGAGCGCTCTCAATGATATCTC	654
Oy	554 GGACAGGACTGTGACCTGAGAGTTGAGTGTGCTTACTAGTTTCCAGACCCCTGGGCTTC	613
Dd	655 GCCAGAGCTGTGAGATCGGAATCGCTGTGATTCCACAGGTTTCAACATTAATATTTC	714
Oy	614 AGAGATTTTTCCTAGAGTGTCTGCGAGAGCGACTCATCATGTACAGCAGCTGCAGTTCC	673
Dd	715 AAGGGAATFAGGGA---TTTAGACAGCTGTCTCATCTCTCTTTGGGCCGATGACATC	771
Oy	674 CCACAGCTGGCTTTGTGTGTGAGCGGTTCAAGCCATTCAAGCTTTTGTACCAAGAACT	733
Dd	772 CAACCTGTGGATTTCTGTGTGAGAGACATGATTAATAATCAAGTCTCTTCAACACGAGACT	831
Oy	734 TTCACAGAAATTAAGTACTCATGAC---CCCAAGATGTGATCGTAATTCACAGGA	790
Dd	832 ACTGGGTCTCAGGCGCCAGGTTAACTGACAGCAAGACAGATCTCTCTTTTGGACTGGG	891
Oy	791 AAAGCATCGACTCTTTTAAACACACGCGCTTGCTAGTTTCTATCATGTGTGTGTGAG	850
Dd	892 ACCGAGTAAAGTGTGTGACCGGAGATCGACAGATGTTTITTAACATGACAAAGCTGG	951
Oy	851 ATCCCATGGCACTGTGTAGAGGTCAAGATCTAAGCCCAAGAGCAATGGCGGCTCAAG	910
Dd	952 AGAAGAGAGCCCGAGTGAGGCGCACAGAGAGAAAGAAAGGCAACAGAGGCGCTGTG	1011

Qy	911	CTTGGACACTGTGATGAGAGGCGGTCTCGAAATGTGAGTAATATGTTAAAG	970
Db	1012	CCCTGMAACACTGTGTGAGATGCTGCGGTGTGGCCAGCTCTTCTGTGGCATGTGGGCCCGCAGC	1071
Qy	971	AAACCATGAGGATGCTGAGAGAAGCTCTACACTCAAGGGGTACATCAGCTATCCCGAAGC	1030
Db	1072	ACGGCATGCAGACGGCTGAGGGGCTCTACAGCAAGCAAGGCTACATCAGCTACCCACGSAAG	1131
Qy	1031	AAACCAACATTTTCCCGAGACACTTAAACCTGACGGGTGTGGTGAACAGCAGACCCCGC	1090
Db	1132	AGACCAACCCACTACCCCTGAGAACTTTGACCTGAAGGGGCTCTGTGGGAGCAGCAGGCAAGC	1191
Qy	1091	ATCCACGCTGGGGGGCCCTTTGCCAGACACTTCTAGACGGGGGTGGTCCACCCCGCAGCA	1150
Db	1192	ACCCCTACTGGGCCGACACCGGTGAAGCGGTTGTTAGCAAGAAAGTATCAACCGCCCGCAGA	1251
Qy	1151	ATGGGAACAAAGTCTGACCAAGCTCACCCCTCCATTTCACCCCAACCAATACCAAA---CA	1207
Db	1252	AAGGCATGAGCGCGGCAACCATCCCCCATCACCCCATGAAGTCTCCCAAGAGGGCG	1311
Qy	1208	ACTTACAGGGAGATGACAGCAGCTGTACGAGATTATTTGTGGCATTTCCTGGCTTGTCT	1267
Db	1312	AATTGGGGGGTACGCGCTGGGGGCTCTGTAGTATCATCAACAGAACACTTCAATCGCCACGG	1371
Qy	1268	GCTCCAGAGATGCTCAGGGGGAGGAGACACAGTGAAGATGACATGCTCAGAGAAAGCT	1327
Db	1372	TCAGCCTAGTACTGCAAGTACTGCCAGACACCATCTCTTACAGATTGGCCCGAGCTCT	1431
Qy	1328	TTTGGCCCATGTGCCCTCATGTATTCGGCCGCAAACTATCTGATGTGATTCATATGATC	1387
Db	1432	TCACCTGTCTCCGGGAAGACCGTCTCTCACCAGGGTTCACGAGAGTTCATGCCCCGTGGCAGA	1491
Qy	1388	ACTGAGT---GACAAGTCTCTCCCTGTCTATGAGCAAGATCCCATCTTACAGCCACAGA	1444
Db	1492	GCGTCCCCCTGGAGAGAGGAGCTGCCCCACTTGCACAGGGGGTGAATGCTTCCCTGTGGGGCG	1551
Qy	1445	CCGTGGAGATGGTGGAGCGGGGAGACACAGCCACCCAAACCTGCTCAGCAGGCCGACCTCA	1504
Db	1552	AGGTGAACATGTGTGAGAGACGACGACGAAACCCACCCGACTACTGTAGAGGAGGCCGAGCTCA	1611
Qy	1505	TTTGCCCTCATGAGAGCAATGGCATTTGATACGGATGCACTTCATGCGGAGCAGCATGAGAGA	1564
Db	1612	TCACGCTCATGTGAGAGCAATGGCATTCGGCAGGATGCGCATCCCTGTGCATATCAACA	1671
Qy	1565	CCATCAAAAGCCCGGATGTACGTGGGGCTCACCCCAAGCAAGGGTTCCTCCCTGGGGCAGC	1624
Db	1672	ACATCTGCCAGGGCAACTATGTCTACGGTGGAGAGCGGGCGCGGCTCAAGGCCCAACCAAC	1731
Qy	1625	TGGGAGATGGGACTTGTGAAGGTTATGTATGATCCATGGGCTATGAATATGTCTAAAGCTTGAC	1684
Db	1732	TGCGATGTGCTGTGTGTGTCAGCGGCTACTATAGATTGATGACAGAGCTGGTGTGCTCCACCA	1791
Qy	1685	TCCGGGCTGAACATGGAAGCTGATCTGAAGCTGATCTGTGATGCAAAAAAGCAAAATTTG	1744
Db	1792	TCCCGCAGTGCAGTGTGAGAGAGCAGCTGAACCTGTGATGCGCCAGGAGGACCGACTACCGCC	1851
Qy	1745	TGCTTCTTAAGGAGCAAGTGCAGAAATCAAGACAGTTC	1785
Db	1852	AGTCTCTGGGCCACACCCCTGGAGCTGTTCAMAGGAGAAGTTC	1892

REFERENCE	FEATURES
1 (bases 1 to 2793)	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1	1 (bases 1 to 2793)
2	2 (bases 1 to 2793)
3	3 (bases 1 to 2793)
4	4 (bases 1 to 2793)
5	5 (bases 1 to 2793)
6	6 (bases 1 to 2793)
7	7 (bases 1 to 2793)
8	8 (bases 1 to 2793)
9	9 (bases 1 to 2793)
10	10 (bases 1 to 2793)
11	11 (bases 1 to 2793)
12	12 (bases 1 to 2793)
13	13 (bases 1 to 2793)
14	14 (bases 1 to 2793)
15	15 (bases 1 to 2793)
16	16 (bases 1 to 2793)
17	17 (bases 1 to 2793)
18	18 (bases 1 to 2793)
19	19 (bases 1 to 2793)
20	20 (bases 1 to 2793)
21	21 (bases 1 to 2793)
22	22 (bases 1 to 2793)
23	23 (bases 1 to 2793)
24	24 (bases 1 to 2793)
25	25 (bases 1 to 2793)
26	26 (bases 1 to 2793)
27	27 (bases 1 to 2793)
28	28 (bases 1 to 2793)
29	29 (bases 1 to 2793)
30	30 (bases 1 to 2793)
31	31 (bases 1 to 2793)
32	32 (bases 1 to 2793)
33	33 (bases 1 to 2793)
34	34 (bases 1 to 2793)
35	35 (bases 1 to 2793)
36	36 (bases 1 to 2793)
37	37 (bases 1 to 2793)
38	38 (bases 1 to 2793)
39	39 (bases 1 to 2793)
40	40 (bases 1 to 2793)
41	41 (bases 1 to 2793)
42	42 (bases 1 to 2793)
43	43 (bases 1 to 2793)
44	44 (bases 1 to 2793)
45	45 (bases 1 to 2793)
46	46 (bases 1 to 2793)
47	47 (bases 1 to 2793)
48	48 (bases 1 to 2793)
49	49 (bases 1 to 2793)
50	50 (bases 1 to 2793)
51	51 (bases 1 to 2793)
52	52 (bases 1 to 2793)
53	53 (bases 1 to 2793)
54	54 (bases 1 to 2793)
55	55 (bases 1 to 2793)
56	56 (bases 1 to 2793)
57	57 (bases 1 to 2793)
58	58 (bases 1 to 2793)
59	59 (bases 1 to 2793)
60	60 (bases 1 to 2793)
61	61 (bases 1 to 2793)
62	62 (bases 1 to 2793)
63	63 (bases 1 to 2793)
64	64 (bases 1 to 2793)
65	65 (bases 1 to 2793)
66	66 (bases 1 to 2793)
67	67 (bases 1 to 2793)
68	68 (bases 1 to 2793)
69	69 (bases 1 to 2793)
70	70 (bases 1 to 2793)
71	71 (bases 1 to 2793)
72	72 (bases 1 to 2793)
73	73 (bases 1 to 2793)
74	74 (bases 1 to 2793)
75	75 (bases 1 to 2793)
76	76 (bases 1 to 2793)
77	77 (bases 1 to 2793)
78	78 (bases 1 to 2793)
79	79 (bases 1 to 2793)
80	80 (bases 1 to 2793)
81	81 (bases 1 to 2793)
82	82 (bases 1 to 2793)
83	83 (bases 1 to 2793)
84	84 (bases 1 to 2793)
85	85 (bases 1 to 2793)
86	86 (bases 1 to 2793)
87	87 (bases 1 to 2793)
88	88 (bases 1 to 2793)
89	89 (bases 1 to 2793)
90	90 (bases 1 to 2793)
91	91 (bases 1 to 2793)
92	92 (bases 1 to 2793)
93	93 (bases 1 to 2793)
94	94 (bases 1 to 2793)
95	95 (bases 1 to 2793)
96	96 (bases 1 to 2793)
97	97 (bases 1 to 2793)
98	98 (bases 1 to 2793)
99	99 (bases 1 to 2793)
100	100 (bases 1 to 2793)
101	101 (bases 1 to 2793)
102	102 (bases 1 to 2793)
103	103 (bases 1 to 2793)
104	104 (bases 1 to 2793)
105	105 (bases 1 to 2793)
106	106 (bases 1 to 2793)
107	107 (bases 1 to 2793)
108	108 (bases 1 to 2793)
109	109 (bases 1 to 2793)
110	110 (bases 1 to 2793)
111	111 (bases 1 to 2793)
112	112 (bases 1 to 2793)
113	113 (bases 1 to 2793)
114	114 (bases 1 to 2793)
115	115 (bases 1 to 2793)
116	116 (bases 1 to 2793)
117	117 (bases 1 to 2793)
118	118 (bases 1 to 2793)
119	119 (bases 1 to 2793)
120	120 (bases 1 to 2793)
121	121 (bases 1 to 2793)
122	122 (bases 1 to 2793)
123	123 (bases 1 to 2793)
124	124 (bases 1 to 2793)
125	125 (bases 1 to 2793)
126	126 (bases 1 to 2793)
127	127 (bases 1 to 2793)
128	128 (bases 1 to 2793)
129	129 (bases 1 to 2793)
130	130 (bases 1 to 2793)
131	131 (bases 1 to 2793)
132	132 (bases 1 to 2793)
133	


```

Db 475 AGGGGAGAACATCTGCTTGGAGGTTCTTGATGCTGTTTCGCCCTCATGTAACAGGCC 534
449 AG-----GTTGGAGCCCGATTCCTGAGATCAGACCCCATGCCCTCA 493
535 ATGTGGCGAAGAACCGTGTTCGGGGCAGGTTTACCTCATCAGCAGACATCACT 594
494 GGACAGCTTGAAAACTCAGACCGCTGATCAGAGGGTGAGCGATGCTGTGATGTA 553
595 GTAATGCCATGGCTGCTAGGCGCCTGACACAGAGGCGCTCTCATGTGATGCTC 654
554 GCGAGAGCTGAGACTGAGATTTGAGACTGCTTACTAGTTCCAGACCTGGGGCTC 613
655 GCGAGAGCTGAGCTGAGATCGGCTGCTGATTCACAGGTTTCAGATAATATTTCC 714
614 AGAGATTTTTCCTGAGGCTGCGACAGAGCTCATCAGTTAGGAGCTGCGAGTTCC 673
715 AGGGGAATACGTGA---TTTACACAGCTCTCTCATCTCTTGGGCGGTGTAGACTC 771
674 CCACACTGGGCTTTGTGTGGAGCGGTTCAAGGCCATTCAGGCTTTGTACCAAAATCT 733
772 CAACCCCTGGATTCTGTGTGAGACATGATATAATCCATCCCTCAACCAAGAGACT 831
734 TCCACAGATTAAGTACATCATGAC---CACAAAGATGGTATGTAATTCACACTGA 790
832 ACTGGGTCTGCGAGCGCAAGGTTACACTGACAAAGACAGATCTCTCTTTGGAGCTGG 891
791 AAAGCATGCACTCTTTAACACACAGGCTTGCCTAGTTCTCTATCAGTTGTGTGGAG 850
892 ACCGATGAAGTGTGTGACCGGAGATCGACAGATGTTTAAACATGACAAAGCTGG 951
851 ATCCCATGCGCACTGTGTGAGGTGACATCTAAGCCCAAGAGCAAGTGGGCTCTAAG 910
952 AGAAGAAAGCCAGGTGAGGCGCAAGAAAGAAAGAAAGGCAACAGAGGCGCCCTGG 1011
911 CCTTGACACTGTGAGCTTGAGAAAGCTGGCTTCTCGAAAGTTGAGATAAATGCTAAG 970
1012 CCTGAAACACTGTGTGAGATGCTGCTGTGGCCAGCTCTTCTGTGGCATGGGGCGCAGC 1071
971 AAACATGAGGATTTGCTGAGAAAGCTTACACTCAAGGATCAGTATCCCGGAACG 1030
1072 ACGCATGACAGCGGCTGAGGGGCTTACACGAAAGGCTACATGACTACCCAGGAGAG 1131
1031 AAACAAACATTTTCCAGAGACTTAAACTGACGGTGTGTGTGGAACAGACAGCCCCG 1090
1132 AGACCAACCCACTACCTGAGAACTTGAAGGCTCTTGTGCGCAGAGAGGCCAAC 1191
1091 ATCCACGCTGGGGGCTTTGCCAGAGCATTTCTAGAGCGGGGTGCCACCCAGCGCA 1150
1192 ACCCTTACTGGCGGACACGCTGAGCGGTTGTTAGCAAGAAAGTATACACCGCCGCGGA 1251
1151 ATGGGAACAAGTGTGACCAAGCTCACCTCCATTCACCCACCAATAACACAA---CA 1207
1252 AAGGCCATGACGCGGCGGACCATCCGCCATACCCCATGAAATGTGCACAGAGCGG 1311
1208 ACTTACAGGAGATGAACACAGACTGACAGTTTATTTGCCATTTCTGGCTTGTCT 1267
1312 AATTAGGGGGTGAAGCGCTGTGCTATAGTACATCACAGACTTATTCGCCACGG 1371
1268 GCTCCCGATGTGTCAGGGGCGAGAGACACAGTGGAGATGACATGCTCAGAGAAAGCT 1327
1372 TCAGCCATGACTGCAAGTACTGCAAGACACATCTCTTACAGAAATGGGCGCAGAGCT 1431
1328 TTGTGGCCATGGGCTCATGATTCCTGCGCGAAATCTATCTGATGTATCATATATATC 1387
1432 TCACCTGTCTCCGGGAAGACGCTCTCTACACAGGCTTACAGAGATGATGCCCTGGAGA 1491
1388 ACTGGAGT---GACAGATCTCTCTGTCTATGAGCAAGATCCCACTTTCAGCCCGACA 1444
1492 GCGTGGCCCTTGAGGAGAGAGCTGCCACTTGGCGGCGGTGATGCTTCTCTGTGGGCG 1551
1445 CCGTGGAGATGATGAGAGGAGAGACGACGACCAAGCTGCTCAGCGAGGCGGAGCTCA 1504

```

```

Db 1552 AGTGAAGATCTGTGAGAGACAGACGAACCAACCCGACTACCTGACGAGGCCGAGCTCA 1611
1505 TTGGCCCTCATGAGAGAGATGGCATTTGGTACGATGCTCATCATCGAGACATCGAGA 1564
1612 TCAGCTCATGAGAGAGATGGCATGGCAGCGAGATGCCAGCATCCCTGTGATATACCA 1671
1565 CCATCAAAAGCCGGATGTAGCTGGGCTCACCCAGACAGAGGTTCTCTCTGGGGACC 1624
1672 ACATCTGCCACCGCAACATATGTACAGGGTGAAGAGGCGGCGGCTCAAGGCCACCAAC 1731
1625 TGGGATGGAGCTTGTGAGAAAGTTATGATTCAGATGGGCTATGAATGTCTAAGCTGACC 1684
1732 TGGCATGCTCTCTGTGCGACGCTACTATTAAGATTGATGACAGAGCTGTCTCCACCA 1791
1685 TCCGGGCTGAACGTGAGAACTGATGAACTGATGCTGATGCAAAAAAGACAATTTG 1744
1792 TCCGAGATGAGTGAAGAGAGAGTGAACCTGATGAGCCAGGCGCAGGCGCATCCGCC 1851
1745 TGGTTCTAAGCAGCAGCAAGTGCAGAAATACAAAGCAGTTTC 1785
1852 AGGTCTGTGGCCACACCTGGAGCGTGTCAAGAGGAGCTTC 1892

RESULT 15
BC002432 3133 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, topoisomerase (DNA) III beta, clone MGC:1867
DEFINITION IMAGE:3346377, mRNA, complete cds.
ACCESSION BC002432
VERSION BC002432.1 GI:12803238
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3133)
REFERENCE 1
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
nisc.mgc@nhgri.nih.gov
Contact: Shevchenko, Y., Wetherby, R.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stanlipop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 5 Row: e Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4102878.
location/Qualifiers
1. 3133
/organism="Homo sapiens"
/db_xref="LocusID:8940"
/db_xref="taxon:9606"
/clone="MGC:1867 IMAGE:3346377"
/tissue_type="Skin, melanotic melanoma."
FEATURES
source

```


Qy 1715 TGATCTGTGATGGCAAAAAGGACAAATTTGTGGTTCTAAGCGCAGCAAGTGCAGAAATACA 1774
 ||||| ||||| || || || || || ||
Db 2153 TGATCGCCCGCAGGCGCAAGCGACTACCGCCAGAGTCCCTGGGCCACACACCTGGACGTGTTCA 2212
Qy 1775 AGCAGGTTTTC 1785
 || ||
Db 2213 AGAGGAAGTTC 2223

Search completed: October 13, 2002, 18:37:48
Job time : 7935.08 secs